Universidad Carlos III de Madrid

# TESIS DOCTORAL

## Título de la tesis:
## Measuring Financial Risk

**Autor**: Maria Rosa Nieto Delfin

**Directora**: Esther Ruiz

Doctorado en Economía de la Empresa y Métodos Cuantitativos
Departamento de Estadística

Universidad Carlos III de Madrid
Getafe, Mayo de 2010

TESIS DOCTORAL

TÍTULO DE LA TESIS:
**Measuring Financial Risk**

Autor: Maria Rosa Nieto Delfin

Directora: Esther Ruiz

Firma del Tribunal Calificador:

| | Nombre y Apellidos | Firma |
|---|---|---|
| Presidente | ROSA ELVIRA LILLO RODRIGUEZ | |
| Vocal | HENRYK GZYL BUCHHOLZ | |
| Vocal | RICARDO CAO ABAD | |
| Vocal | ANGEL LEON VALLE | |
| Secretario | ANDREAS HEINEN | |

Calificación:

Getafe, de de 2010

*To my Grandmother, my Aunt and my Dad*

# Contents

**Acknowledgements** **5**

**Resumen** **6**

**1 Introduction** **9**

1.1 Motivation . . . 9

1.2 Value at Risk and Expected Shortfall . . . 14

1.2.1 Alternative estimators . . . 14

1.2.2 Alternative horizons and levels . . . 15

1.2.3 Computing the uncertainty of VaR and ES . . . 17

1.3 Organization of the thesis . . . 18

**2 Measuring Financial Risk: Comparison of alternative procedures to estimate VaR and ES** **20**

2.1 Introduction . . . 20

2.2 Estimation and testing of Value at Risk ($VaR$) . . . 21

2.2.1 Estimation Methods for $VaR$ . . . 21

2.2.2 Backtesting VaR estimates . . . 31

2.3 Estimation and testing of Expected Shortfall ($ES$) . . . 35

2.3.1 Estimation . . . 35

2.3.2 Backtesting ES estimates . . . 37

2.4 Empirical Application: Estimating the $VaR$ and $ES$ of S&P500 returns. 38

2.4.1 Model fitting . . . 38

2.4.2 Estimates of the $VaR$ . . . 40

2.4.3 Estimates of the $ES$ . . . 43

2.5 Conclusions . . . 44

**3 Robustness against VaR and ES level and horizon** **67**

3.1 Introduction . . . . . . . . . . 67
3.2 Effects of different VaR and ES levels . . . . . . . . . . 68
3.3 Alternative VaR and ES horizons . . . . . . . . . . 71
3.4 Conclusions . . . . . . . . . . 75

**4 Bootstrap Prediction Intervals for Risk Measures in the context of GARCH models 101**

4.1 **Introduction** . . . . . . . . . . 101
4.2 Bootstrap prediction intervals for VaR and ES . . . . . . . . . . 102
4.2.1 Bootstrap based prediction intervals for VaR and ES . . . . . . 103
4.2.2 A new bootstrap procedure . . . . . . . . . . 109
4.3 Monte Carlo experiments . . . . . . . . . . 111
4.4 Empirical Application . . . . . . . . . . 114
4.5 Conclusions . . . . . . . . . . 115

**5 Summary of Conclusions and Future Research 121**

**Bibliography 124**

# Acknowledgements

First of all, I would like to thank my supervisor Esther Ruiz for her patience, understanding, help and teaching along these last years. Fortunately, I learned a lot from her, not only about the subject of this thesis, but also as a person. I really appreciate all the time she spent working with me, because without her support I would have not finished this work.

I'm very grateful to the professors of the Statistics Department for their support and help with all the issues related to the teaching practice, specially with the coordination assistantship.

I would also like to show my gratitude to the Spanish Ministry of Education and Science and to the education and culture section of the Community of Madrid for financial support. In particular, I would like to thank Esther Ruiz for allowing me to participate in the research projects SEJ2006-03919 and ECO2009-08100.

As I promised, I will specially mention two of the most important people I knew in Spain, my friends Santi and Alex. I really do not have words to thank them for all the support and advices they gave me and also for the time they dedicated to help me with uncountable problems. I always said, as a joke, that they were my "co-advisors". Thanks a lot, I know we will be friends forever.

I also want to thank the rest of my friends for being my family here. Ana, Alba, Pepa, Ester, Ana Laura, Peter, Adolfo, Andre, Hugo and Silviu. Thanks for the talks, the moments, the anecdotes and specially for the company. You can always rely on me.

I dedicate my thesis to the three most important columns of my life, my grandmother Rosa Montero, my aunt Rosa A. Delfin and my father Isaac Arturo Nieto, because they though me all the fundamental lessons that I keep in my heart and that I try to apply everyday that I live. I will also want to thank my sisters Adriana, Teresa and Larissa and my brother Isaac Antonio for standing next to me along all these years. Thanks for taking care of me and for your advices. My very special mention is for the fundamental part of my life, my six kids. My nieces Angie, Tazi, Yoryis and Sarita and my nephews Nemo and Puny, you are the reason why my life is happier now. Since you were born, everyday is different and full of joy. I am so grateful of having you with me and I am happy that very soon I will be around for watching you grow.

# Resumen

La importancia de la administración de riesgos viene de la necesidad que tienen los bancos y entidades financieras de tener una reserva de capital que les permita afrontar sus obligaciones financieras. El concepto de riesgo es muy amplio debido a que hay diferentes grupos de personas interesados en la bolsa de valores y cada grupo tiene su propia actitud con respecto al riesgo; ver Granger (2002). El riesgo financiero puede ser clasificado en riesgo de crédito, de liquidez, operacional, legal y de mercado. Riesgo de crédito es el riesgo que se adquiere cuando las contrapartes no son capaces de cumplir con sus obligaciones contractuales. Riesgo de liquidez es la inhabilidad para efectuar pagos contraídos con anterioridad. El riesgo operacional está relacionado con accidentes técnicos y humanos, el riesgo legal surge cuando una transacción no puede ser legalmente completada. Finalmente, el riesgo de mercado es el riesgo asociado con cambios inesperados en los rendimientos en intervalos cortos de tiempo. En esta tesis nos centraremos en el riesgo de mercado; ver Jorion (1990) y Duffie and Pan (1997).

Existen dos problemas importantes cuando se trata de estimar el riesgo. Primero, se deben considerar medidas de riesgo con propiedades teóricas adecuadas. Segundo, se deben escoger estimadores con propiedades estadísticas apropiadas.

Una de las medidas de riesgo más populares es el Valor en Riesgo (VaR). El VaR aparece como consecuencia de algunos resultados adversos a lo largo de la historia que forzaron a las agencias reguladoras de la actividad financiera a buscar una forma cuantitativa de definir el riesgo asociado a una posición en el mercado. El VaR se define como la mínima pérdida potencial que, en el $100\alpha\%$ de los peores casos con $\alpha \in (0, 1)$, puede tener una cartera en un horizonte temporal determinado. Entre las principales ventajas del VaR están su simplicidad, aplicabilidad y universalidad; ver Jorion (1990,

1997) y Embrechts et al. (2000). Sin embargo, tiene importantes limitaciones desde el punto de vista teórico. El inconveniente mas importante de esta medida, es que el VaR de una cartera diversificada puede ser mayor que la suma de los riesgos de las carteras individuales.

Como resultado de las limitaciones del VaR como medida de riesgo, Artzner et al. (1997) definieron lo que se conoce como Medidas de Riesgo Coherente. Artzner et al. (1999) propusieron el Tail Conditional Expectation o también llamado Conditional Value at Risk (CVaR). El CVaR mide la pérdida esperada en que se incurrirá en el $100\alpha\%$ de los peores casos. Adicionalmente, Acerbi and Tasche (2002) proponen el Expected Shortfall (ES) como medida de riesgo coherente. Es importante mencionar que el ES es igual al CVaR cuando la distribución de los rendimientos es continua.

Sin embargo, el VaR sigue siendo la medida mas utilizada por los bancos e instituciones financieras. Además, una adecuada estimacin del VaR es fundamental para estimar el ES. Por lo tanto, existe un gran interés en su estimación. Hay varios temas relacionados con la estimación del VaR y del ES que serán considerados en esta tesis. Primero, la decisión acerca del estimador que se utilizará. Segundo, se necesita escoger el nivel $\alpha$ para el VaR y el ES así como el periodo sobre el cual se calcularán ambas medidas. Finalmente, es también importante tener medidas sobre la incertidumbre asociada con la estimación.

En el Capítulo 2 se revisan varios estimadores para el VaR y el ES. Las ventajas y desventajas de dichos estimadores son ilustradas implementándolos a los rendimientos diarios del $S\&P500$. También se revisan y comparan mtodos alternativos para probar la precisión de las estimaciones del VaR y del ES. El objetivo en este Capítulo es describir las principales contribuciones en estimación de ambas medidas de riesgo actualizando estudios previos. Además, extendemos estos estudios con una comparacin de métodos mas exhaustiva. Se consideran varios modelos alternativos para la varianza condicional y para la distribución de los errores. Finalmente, también se comparan algunos estimadores propuestos en la literatura para estimar el ES.

En el Capítulo 3 se analizan los resultados que se obtienen cuando, en lugar del requerido 1%, se consideran puntos diferentes de la cola de la distribución de los rendimientos, por ejemplo, el 5% y el 10%. Se implementan los procedimientos de estimación descritos en el Capítulo 2 y se comparan los resultados con los que se obtenían al 1%. Adicionalmente, se analizan los procedimientos utilizados para predecir el VaR y el ES en horizontes de predicción distintos a un periodo hacia adelante. El comité de Basilea requiere que el VaR sea reportado en periodo de 10 das. Por esta razón, el análisis se ha enfocado en predecir en este horizonte. Se han implementado y comparado distintos procedimientos a la serie de rendimientos diarios y quincenales del $S\&P500$.

Finalmente, en el Capítulo 4 se toma en cuenta la incertidumbre asociada con la estimación del VaR y del ES mediante la construcción de intervalos de predicción. Se supone que los rendimientos están bien representados por modelos de tipo GARCH y se propone una extensin del procedimiento bootstrap de Christoffersen and GonÇalves (2005) mediante la incorporación de un segundo paso bootstrap en la estimación del percentil de la distribución condicional de los residuos estandarizados. Además, siguiendo la propuesta de Ho and Lee (2005), se consideran intervalos de predicción bootstrap que superan las limitaciones de los intervalos de predicción tradicionales. Se muestra que nuestro procedimiento bootstrap mejora el rendimiento de los intervalos de predicción para el VaR y el ES al tener coberturas más cercanas a las nominales.

# Chapter 1

# Introduction

## 1.1 Motivation

The importance for risk management in financial institutions comes from the necessity of having a reserve of capital for facing their financial obligations. The concept of financial risk is very wide since there are different groups of people interested in the money market and each group has its own attitude about risk; see Granger (2002). Financial risk can be classified into credit, liquidity, operational, legal and market risks. Credit risk is the risk faced when the counterparties are unable to fulfill their contractual obligations while liquidity risk refers to the inability to meet payments obligations, operational risk is related to human and technical accidents and legal risk arises when a transaction cannot be legally accomplished. Finally, market risk is the risk associated with unexpected changes in returns over short time horizons; see Jorion (1990) and Duffie and Pan (1997). In this thesis, we focus on market risk. In a simple situation, if we buy an asset at price $P_{t-1}$ at time $t-1$, and sold it at price $P_t$ at time $t$, we get a return calculated as the first difference of logarithm of prices, $R_t = \log(P_t) - \log(P_{t-1})$. In $t-1$, $R_t$ is unknown and the risk is caused by this uncertainty. The return at time $t$ can be considered unsatisfactory by the investor when it is negative or inferior to the return of some kind of governmental bond.

There are two main issues involved in estimating risk. First, one should consider measures of risk with adequate properties from a theoretical point of view. Second, once we decide how to measure risk, we should choose estimators of the corresponding measure with appropriate statistical properties.

There are different measures of market risk proposed in the literature. Luce (1981) suggests to measure risk by assigning different weights to the two halves of the distribution of returns. Therefore, if $f(R)$ is the density of returns, the risk is given by

$$Risk_t = A_1 \int_0^{\infty} R_t^{\theta} f(R)\, dR_t + A_2 \int_{-\infty}^{0} |R_t|^{\theta} f(R)\, dR_t \tag{1.1}$$

where $A_1, A_2 \geq 0$ and $\theta > 0$. Depending on whether $f(R)$ is the marginal density or the density of $R_t$ conditional on past observations, we obtain marginal or conditional moments. In this thesis, we consider conditional distributions of returns because it is a more efficient use of the information contained on the data. When the weights in (1.1) are equal, we have the class of volatility measures given by

$$V_t(\theta) = \underset{t-1}{E}\left[|R_t - \mu_t|^{\theta}\right] \tag{1.2}$$

where $\mu_t = \underset{t-1}{E}[R_t]$ and the $t-1$ under the expectation means that it is taken conditional on the information available up to time $t-1$. $V_t(\theta)$ includes the two most popular measures of risk, namely the variance, when $\theta = 2$, and the mean absolute deviation, when $\theta = 1$. However, only when the utility function is quadratic or the distribution of returns is Normal or log-Normal, the variance is an appropriate measure; see Tobin (1969), Tsiang (1972), Machina and Rothschild (1987) and Levy (1992). The assumption of Normal conditional distribution could be adequate in some financial returns. However, the utility cannot be assumed to be quadratic as the investor has different attitudes depending on whether the returns are over or under their means. In that sense, there is uncertainty in the upper part of the distribution, but the risk only exists in the lower part, which means that the investors do not diversify for reducing the possibility of an unexpected positive return, just if it is negative; see Granger (2002). Therefore, measures based on $V_t(\theta)$ are not in general adequate to measure risk.

One of the most popular alternative measures of risk is what is known as the Value at Risk ($VaR$). The $VaR$ appears as a consequence of some adverse results along history which force the agencies that regulate financial activity to look for a quantitative way to define the risk associated to a position in the market. The $VaR$ is defined as the minimal potential loss that a portfolio can suffer in the $100\alpha\%$ worst cases with $\alpha \in (0,1)$, on some fixed time horizon. In particular, the $VaR$ is given by

$$VaR_t^{\alpha} = \sup \left[ r \mid \underset{t-1}{P} \left[ R_t \leq r \right] \leq \alpha \right]. \tag{1.3}$$

Among the main advantages of the $VaR$ are simplicity, wide applicability and universality; see Jorion (1990, 1997), and Embrechts et al. (2000). Consequently, since the 80's, the regulatory agencies have used the $VaR$ to measure the risk of financial institutions. According to the Basel Committee on Banking Supervision, banks are required when calculating $VaR$. They require to compute the $VaR$ for $\alpha = 0.01$ and for returns corresponding to 10 trading days. Furthermore, the $VaR$ should be computed with observations corresponding to at least one year. However, the $VaR$ has fundamental limitations from the point of view of its theoretical properties. The most important of them is that the $VaR$ of a diversified portfolio can be greater than the sum of the $VaRs$ of the individual portfolios; see Acerbi and Tasche (2002). Furthermore, the $VaR$ does not measure losses exceeding itself. Consequently, we can have two distributions with heavy tails and the same $VaR$, but the losses that exceed $VaR$ could be totally different; see Basak and Shapiro (2001) and Yamai and Yoshiba (2005) for both theoretical and practical discussions on the tail risk of $VaR$. In order to exemplify this, Acerbi et al. (2001) used the next paradox: consider a portfolio A (made of long positions) of value 1000 euro with a maximum downside level of 100 euro and suppose that the worst 5% cases on a fixed time horizon T are all of maximum downside. $VaR$ at 5% on this time horizon would then be 100 euro. Consider now another portfolio B again of 1000 euro which on the other hand invests also in strong futures positions that allow for a potential unbounded maximum loss. We could choose B in such a way that its $VaR$ is still 100 euro on the time horizon T. However, in portfolio A the 5% worst case losses are all of 100 euro and in portfolio B the 5% worst case losses range from 100 euro to some arbitrarily high value. Additionally, from the point of view of optimization, the $VaR$ is not useful because it is not convex; see Szegö (2002).

As a result of the limitations of the $VaR$ as a measure of risk, Artzner et al. (1997) define what is known as Coherent Measures of Risk. A Coherent Measure of Risk, $\rho(\cdot)$, must satisfy the following properties:

**(i)** Monotonicity: $\forall\ R, S$ returns of two assets of portfolios, such that $R \leq S \Longrightarrow \rho(R) \leq \rho(S)$;

**(ii)** Positive Homogeneity: $\forall c \geq 0$ and $\forall R$, $\rho(cR) = c\rho(R)$;

**(iii)** Translation invariance: $\forall r \in \mathbb{R}$ and $\forall R$, $\rho(R + r) = \rho(R) - r$;

**(iv)** Subadditivity: $\forall\ R, S$, $\rho(R + S) \leq \rho(R) + \rho(S)$.

The monotonicity condition implies that if the return of a portfolio is smaller than that one of another portfolio, then the portfolio with larger return will have larger risk. The second property means that if the return is multiplied by a constant, the risk will change in the same proportion. The third property means that if you invest in a risk free asset, the faced risk will decrease in this amount. The most distinctive of these properties is the Sub-additivity which implies that a portfolio which is made of sub-portfolios would have at most the same risk as the sum of the risks of sub-portfolios thanks to risk diversification. Note that, as we mentioned above, the $VaR$ is not sub-additive.

Wang et al. (1997) propose measures of risk with distortion functions which, under certain conditions, are coherent; see Wang (1998) for one of this measures. Later, Artzner et al. (1999) propose the Tail Conditional Expectation or Conditional Value at Risk ($CVaR$). The $CVaR$, that measures the *expected loss* in the $100\alpha\%$ worst cases, is given by

$$CVaR_t^\alpha = \underset{t-1}{E}\left\{R_t | R_t \leq -VaR_t^\alpha)\right\}. \tag{1.4}$$

The $CVaR$ is a coherent measure of risk when it is restricted to continuous distributions. However, it can violate sub-additivity with non-continuous distributions. Consequently, Acerbi and Tasche (2002) propose the Expected Shortfall ($ES$) which is given by

$$ES_t^\alpha = -(CVaR_t^\alpha + (\lambda - 1)\left(CVaR_t^\alpha - VaR_t^\alpha\right)) \tag{1.5}$$

where $\lambda \equiv \frac{\underset{t-1}{P}\left[R_t \leq -VaR_t^\alpha\right]}{\alpha} \geq 1$. Note that $CVaR = ES$ when the distribution of returns is continuous; see Giannopoulos and Tunaru (2005). However, the $ES$ is still coherent when the distribution of returns is not continuous. Another advantage of the $ES$ when compared with the more popular $VaR$, is that it is free of tail risk in the sense that it takes into account information about the tail of the underlying distribution. The use of free tail risk measures avoids extreme loss in the tail. Therefore, the $ES$ is an excellent candidate for replacing $VaR$ for financial risk management purposes. However, the effectiveness of $ES$ depends on the stability of its estimation and the choice of efficient backtesting methods; see Fabozzi and Tunaru (2006).

Although the $VaR$ has important theoretical limitations as a measure of risk, Danielsson et al. (2005) explore the potential for violations of the $VaR$ subadditivity and conclude that for most practical applications the $VaR$ is subadditive. Therefore, according to this analysis, there is no reason to choose a more complicate risk measure than the $VaR$ solely for reasons of coherence. Furthermore, Danielsson et al. (2006) show that for heavy tailed distributions, as those observed in financial returns, the choice of downside risk measures does not seem to matter much as all of them (including $VaR$ and $ES$) order heavy tailed risk in a similar manner. In any case, the $VaR$ is still the measure most extensively implemented by banks and financial institutions. Furthermore, an adequate estimation of the $VaR$ is fundamental to estimate the $ES$. Therefore, there is a huge interest on its estimation. There are several issues of interest related with the estimation of the $VaR$ and $ES$ that will be considered in this thesis. First, one has to decide about the particular estimator to be implemented. Second, the level of the $VaR$ and the $ES$ has to be chosen as well as the period for computing them. Finally, it is also important to have measures of the uncertainty associated with their estimation.

# 1.2 Value at Risk and Expected Shortfall

## 1.2.1 Alternative estimators

Once one decides about which measure of risk to implement, it is necessary to estimate it. Although the Basel Committee establishes that financial institutions must use the $VaR$ as a measure of risk, they are free to choose the estimation method. There are a very large number of estimation methods proposed in the literature to estimate the $VaR$ and a relative smaller number of proposals for the $ES$; see Manganelli and Engle (2001), GenÇay and SelÇuk (2004), Angelidis et al. (2005), Kuester et al. (2006), Lima and Néri (2007) and McAleer and da Veiga (2008) for surveys on $VaR$ and $ES$ estimation.

$VaR$ estimators can be classified in nonparametric, semiparametric and parametric. Some of the most popular nonparametric methods are Historical Simulation ($HS$) and Bootstrap procedures that do not make any distributional assumption on the distribution of returns. On the other hand, there are many proposals in the literature of parametric specifications of the conditional mean and variance and the conditional error distribution implemented to estimate the VaR and ES, as for example, the Riskmetrics, $GARCH$ or $CAViaR$ specification. Finally, some examples of the semiparametric methods are those based on Extreme Value Theory ($EVT$) and Feasible Historical Simulation ($FHS$). Both methods assume $GARCH$ models for the conditional volatility but differ in the way the quantile of the distribution of returns is calculated.

After the $VaR$ and $ES$ are estimated, one needs to measure their accuracy. The Basel Committee proposes a backtesting procedure which consists on the comparison between the nominal $VaR$ level, $\alpha$, and the proportion of actual returns which are less than or equal to the $VaR$ forecasts. Many backtesting procedures have been proposed afterwards in the literature; see, for example, Kupiec (1995), Christoffersen (1998), Christoffersen and Diebold (2000), Christoffersen et al. (2001), Dowd (2001), Engle and Manganelli (2004) and Berkowitz et al. (2006) for the $VaR$ and Berkowitz (2001) and Kerkhof and Melenberg (2002) for $ES$. These backtesting procedures are designed to discriminate whether a particular estimator is accurate. However, when several alternative accurate estimators are available, one also wants to choose which is the best among them. There are several proposals to accomplish this task, as for example, Lopez (1999), Sarma et al. (2003), Giacomini and Komunjer (2005), Bao et al. (2006), Giacomini and White (2006) and Angelidis and Degiannakis (2007).

Very recently, Wong (2010) proposes the tail risk statistic for backtesting. This statistic provides information on the risk faced by investors beyond the $VaR$ boundary. Note that the tail risk statistics is closely related with the $VaR$ and ES but is an alternative measure of risk which we will not consider further in this thesis.

Our first goal is to make an updated and detailed revision of the literature on estimation and backtesting of $VaR$ and $ES$. By implementing alternative procedures to the same series of returns, we will analyze whether there are significant differences between the estimates obtained by alternative procedures and, in this case, whether it is more important to have an appropriate specification of the conditional variance or the distribution of returns.

### 1.2.2 Alternative horizons and levels

In practice, when a risk manager faces the problem of measuring the risk of a portfolio by estimating the $VaR$ or the $ES$, he must take into account that the final conclusion will change depending on the $\alpha$ level, the horizon and the sample period of the data. Therefore, these three factors must be fixed at the beginning of the analysis depending on the necessities of the financial institution.

Most of the literature related with the estimation of the $VaR$ and $ES$ deals with one-step ahead predictions. However, according to the amendments of the Basel Committee, the $VaR$ should be reported for a 10 days horizon, in such a way that the portfolio manager has time to rebalance his portfolios in case it is needed. Furthermore, there are situations in which measurements of risk are required for longer horizons, for example, financial institutions with long term liabilities like pension funds and life insurance companies, corporations with planning horizons longer than one year, banks when decide on long run policy issues such as economic capital; see Giannopoulos (2003).

There is a rule commonly applied in order to transform one-step ahead $VaR$ to a longer horizon $VaR$, known as the square root of time, which is given by

$$VaR^{\alpha}_{t+h} = \sqrt{h} VaR^{\alpha}_{t+1}. \tag{1.6}$$

However, Kupiec (1995) and Blake et al. (2000) show that, under non-normality, aggregating daily $VaR$ with formula (1.6) is unreliable and can lead to considerable overestimates of the $VaR$; see also Diebold et al. (1997) and Danielsson and Zigrand (2006) with respect to forecasting long-horizon volatility when the time scaling fails for many processes such as $GARCH$, Stochastic Volatility and jump processes. Several alternatives to the square root rule have been proposed in the literature. For example, Historical Simulation is one of the procedures preferred by the risk managers in estimating the $VaR$ for longer horizons. Alternatively, Danielsson and Hartmann (1998), and McNeil and Frey (2000) consider procedures based on $EVT$. While Ruiz and Pascual (2002b) and Giannopoulos (2003) implement bootstrap procedures for $VaR$ estimation in horizons longer than one day. As far as we know, there are not results on forecasting the $ES$ for longer horizons. Only McNeil and Frey (2000) mention that their procedure could be applied in this case without implementing it. We want to analyze if the results obtained for the one-step ahead $VaR$ and $ES$ can be generalized to the ten-steps ahead and also compare the results when using fortnightly data instead of daily data.

On the other hand, some papers in the literature have implemented different methods for estimating the $VaR$ and the $ES$ at the 5% or the 10% level; see Danielsson and de Vries (2000), Haas and Kondratyev (2000), McNeil and Frey (2000), Giot and Laurent (2003), Angelidis et al. (2005), Giannopoulos and Tunaru (2005), Harmantzis et al. (2006), Kuester et al. (2006), Martins-Filho and Yao (2006), Bali and Theodossiou (2007) and Jalal and Rockinger (2008). One of the reasons of considering these confidence levels is because the popular RiskMetrics focuses on the 5% quantile and it is usually compared with the new proposals. Therefore, we implement the procedures used for the estimation of the 1% $VaR$ and $ES$ to these two levels and compare the results.

### 1.2.3 Computing the uncertainty of VaR and ES

When estimating the $VaR$ and the $ES$ it could be important to measure the uncertainty associated with their estimates. It can be useful to report not only point estimates of future $VaR$ and $ES$ but also their corresponding confidence intervals for setting capital requirements and establishing limits for banks and traders. Christoffersen and GonÇalves (2005) provide the next example to illustrate that in practical situations, the interval estimation gives more information than only the point estimation. Assume that a portfolio manager has to construct a portfolio with a $VaR$ up to 15% of the current capital. If he has a point estimate for the $VaR$ of 13% and a confidence band of $10\% - 16\%$ then he should rebalance the portfolio in order to reduce risk. If the point estimate were the only information available, the decision would be that the portfolio is safe and there is no need to rebalance.

There is a large literature devoted to point forecast, theoretical properties and backtesting of $VaR$ and $ES$; see Nieto and Ruiz (2009) for a recent survey in the context of univariate time series of returns. However, there are quite few papers considering prediction intervals for these quantities. For example, Chan et al. (2007) propose to construct confidence intervals for the $VaR$ by the tilting method of Hall and Yao (2003) and Peng and Qi (2003). Chen and Tang (2005a) propose a nonparametric estimation of the $VaR$ and its associated standard error. Chou et al. (2008) and Gilli and Kllezi (2006) construct confidence intervals for the $VaR$ and the $ES$ respectively, using Extreme Value Theory ($EVT$). Finally, Lan et al. (2008) use the statistical theory of empirical likelihood to construct confidence intervals for the $ES$. However, these prediction intervals do not incorporate the uncertainty due to parameter estimation. Bams et al. (2005) shows that incorporating the parameter uncertainty within the prediction intervals for $VaR$ and $ES$ is important. Consequently, they use the asymptotic covariance matrix of the Maximum Likelihood ($ML$) estimator to quantify the uncertainty of the $VaR$ by sampling from the asymptotic parameter distribution. However, this distribution can be an inadequate approximation of the finite sample distribution when the sample size is small.

Alternatively, it is possible to incorporate the parameter uncertainty by using bootstrap procedures which work well in prediction; see, for example, the survey by Ruiz and Pascual (2002a). In that sense, Christoffersen and GonÇalves (2005) propose using bootstrap procedures to obtain prediction intervals for several parametric and nonparametric estimates of the $VaR$ and $ES$. In the case of the parametric estimates, they consider a univariate $GARCH(1,1)$ model for the conditional variances and implement the bootstrap procedure of Pascual et al. (2006). Then, in order to compute the corresponding quantile needed for the prediction of the $VaR$ and $ES$, they consider several alternative assumptions about the distribution of the standardized returns. First, they consider a Normal and Student-$\nu$ distributions. Second, they assume an Extreme Value distribution and compute the corresponding quantile by using the Hill estimator. Third, they approximate the distribution using the Cornish-Fisher and Gram-Charlier approximations. Finally, they implement Feasible Historical Simulation ($FHS$). When considering nonparametric estimates of the $VaR$ and $ES$, Christoffersen and GonÇalves (2005) focus on the iid bootstrap procedure to obtain prediction intervals for the $VaR$ and $ES$ computed using Historical Simulation ($HS$). This bootstrap procedure is completely non-parametric avoiding any distributional assumption on the data. However, by implicitly assuming that returns are iid, this method fails to capture the dependence in returns when it exist. Within this context, they conclude that their bootstrap procedure has adequate coverage when the $FHS$ is implemented to estimate the $VaR$. On the other hand, the Hill estimator has the best coverage for the $ES$ but still well under the nominal. It is important to note that from a conservative risk management perspective under-coverage is worst than over-coverage.

## 1.3 Organization of the thesis

The rest of this thesis is organized as follows. Chapter 2 surveys the estimation methods for $VaR$ and $ES$ and the backtesting procedures proposed to measure the accuracy and for selecting the procedure which delivers better one-step ahead forecasts of both risk measures. We describe the characteristics, advantages and disadvantages of each method. The results are illustrated by estimating the $VaR$ and $ES$ of a financial time series of returns. Furthermore, by comparing the $VaR$ and $ES$ estimated with alternative specifications of the conditional variance and error distributions, we show that the former is more important than the latter in order to obtain appropriate estimates of the risk measures considered.

On Chapter 3, we consider the influence on the conclusions of Chapter 2 of computing ten-steps ahead instead of one-step ahead $VaR$ and $ES$ forecasts. We compute them in a financial series of returns with daily and fortnightly observations and compare the results. In this Chapter we also consider different levels for the $VaR$ and $ES$ and show that depending on $\alpha$, the conclusions about the most adequate procedure are not different.

In Chapter 4, we turn our attention to interval estimation of risk. We propose a new bootstrap procedure that extends that proposed by Christoffersen and GonÇalves (2005). This proposed procedure is based on a second bootstrap step from the original residuals instead of using the bootstrap residuals. Consequently, our procedure avoid the estimation error involved in the residuals. We also make an extension of the procedure of Pascual et al. (2006) applied in Ruiz and Pascual (2002b) for $VaR$ estimation, to the case of the $ES$. We carry out Monte Carlo experiments in order to analyze the finite sample properties of our procedure and compare them with those of the alternative procedures as those proposed by Christoffersen and GonÇalves (2005) and Ho and Lee (2005). The latter produces better results in terms of coverage but it has the drawback of the selection of an optimal smoothing bandwidth.

Finally, Chapter 5 summarizes the main conclusions of this thesis and presents some suggestions for future research.

# Chapter 2

# Measuring Financial Risk: Comparison of alternative procedures to estimate VaR and ES

## 2.1 Introduction

In this Chapter, we review several alternative estimators of $VaR$ and $ES$. The advantages and disadvantages of the estimators considered are illustrated by implementing them to the estimation of the $VaR$ and $ES$ of a time series of daily $S\&P500$ returns. We also revise and compare alternative methods to test for the adequacy of $VaR$ and $ES$. The literature on the estimation of the $VaR$ is so large that it is unfeasible trying to cover all the available contributions. Consequently, our objective in this Chapter is to describe the main contributions updating other previous surveys and comparisons. Furthermore, we extend these surveys by providing a more comprehensive comparison of methods. We consider a larger number of: i) models for the conditional variance and ii) error distributions. Finally, we also compare several estimators proposed in the literature to estimate the $ES$.

This Chapter has been organized as follows. Section 2.2 describes several estimation methods for $VaR$ and backtesting procedures to measure its adequacy. Section 2.3 is devoted to reviewing the estimation and backtesting methods for $ES$. Section 2.4 illustrates the estimation methods described in the two previous sections by implementing them to estimate the $VaR$ and $ES$ of a series of daily S&P500 index returns. Additionally, these procedures are compared through backtesting. Finally, Section 2.5

concludes the Chapter with the main conclusions and suggestions for further research.

## 2.2 Estimation and testing of Value at Risk ($VaR$)

This section describes some of the most popular methods to estimating the $VaR$, focusing on the weakness and strengths of each of them. The estimation of $VaR$ is a difficult computational task due to, among other reasons, the complexity of financial instruments, the dimension of portfolio, the assessment of market probabilities, the approximations introduced to speed up computations and the statistical error on its estimation; see Ju and Pearson (1999), Acerbi et al. (2001), Longin (2001), Krause (2003), and Bao and Ullah (2004) among others. When measuring the risk of a portfolio, this portfolio can be considered as a multivariate system of individual returns or as a univariate return of the whole portfolio. In this Chapter, we focus on the estimation of the $VaR$ of a univariate series of returns; see Santos et al. (2009) for an application of multivariate estimation of $VaR$. Additionally, in this section, we describe some backtesting methods used to evaluate the performance of the $VaR$ estimates.

### 2.2.1 Estimation Methods for $VaR$

The oldest and still very popular estimator of the $VaR$ is based on Historical Simulation ($HS$). The $VaR$ is estimated as the $\alpha th$ quantile of the empirical distribution of losses, $\widehat{VaR}_t^\alpha = -R_{\omega:T}$, where $R_{\omega:T}$ is the $\omega th$-order statistic of the data, $\omega = [T\alpha] = \max\{m \mid m \leq T\alpha, m \in \mathbb{N}\}$ and $T$ is the sample size; see Acerbi and Tasche (2002). $HS$ is simple and does not assume any particular distribution of returns. However, it is based on assuming that returns are $iid$ which is an empirically inadequate assumption. Furthermore, it is well known, that empirical quantiles are not efficient estimators of extreme quantiles. In spite of these limitations, several authors conclude that, in practice, $HS$ could generate adequate estimates of the $VaR$ depending on the length of the data and the $VaR$ level, $\alpha$; see, for example, Hendricks (1996) and Vlaar (2000) who obtains satisfactory results when $T = 2,550$ and $\alpha = 0.05$[1].

[1]Remember that the Basel Committee requires the supervision of the VaR for $\alpha = 0.01$.

Another popular estimator of the $VaR$ based on the $iid$ assumption is based on boostrapping. To compute the $VaR$, B series of bootstrap returns $R^* = (R_1^*, ..., R_T^*)$, are drawn with replacement from the original series of returns, with each return having the same probability of being chosen. Then, the $\alpha th$ empirical quantile of each of the $B$ replicates is calculated as in $HS$. Finally, the $VaR$ is estimated as the average of these $\alpha th$ empirical quantiles; see Barone-Adesi and Giannopoulos (2001) for an illustrative example. Note that using this procedure, it is possible to obtain confidence intervals for the estimated $VaR$; see Christoffersen and GonÇalves (2005). However, given that the iid assumption is not appropriate, the properties of the bootstrap procedure are not standard.

Given that the $iid$ assumption is not adequate for real daily returns, there are many alternative estimators based on assuming particular specifications for the conditional distribution of returns. Consider the following model of returns

$$R_t = \mu_t + \epsilon_t \sigma_t \tag{2.1}$$

where $\mu_t$ and $\sigma_t$ are the conditional mean and the conditional standard deviation of returns respectively, and $\{\epsilon_t\}$ are $iid$ disturbances with zero mean and variance 1. Thus, the $100\alpha\%$ one-step ahead $VaR$ conditional on information available at time $t-1$ is given by

$$VaR_t^{\alpha} = \mu_t + q_{\alpha} \sigma_t \tag{2.2}$$

where $q_\alpha$ is the $100\alpha\%$ quantile of $f(\epsilon_t)$, the density of the centered and standardized returns, $\epsilon_t$.

In order to estimate the $VaR$ in (2.1) one needs to specify and estimate the conditional mean and the conditional variance of returns and to assume a particular distribution for $\epsilon_t$. Table 2.1 contains a summary of different assumptions on $\mu_t$, $\sigma_t$ and the distribution of $\epsilon_t$ often made in the literature. The first conclusion from this table is that the most popular assumption for the conditional mean of returns is to specify it as an $ARMA(p, q)$ model given by

$$\mu_t = \phi_0 + \sum_{i=1}^{p} \phi_i R_{t-i} - \sum_{j=1}^{q} \theta_j a_{t-j} \tag{2.3}$$

where $a_t = R_t - \mu_t = \sigma_t \epsilon_t$; see McNeil and Frey (2000), Bali and Theodossiou (2007) and Kuester et al. (2006) among others. Furthermore, given that the dependence on the conditional mean of returns is usually very simple, most authors have represented it by $AR(1)$ or $MA(1)$ models. On the other hand, looking at the specifications of the conditional variance, Table 2.1 shows that many authors choose models within the $GARCH$ family. The simplest of these models is the $GARCH\,(1,1)$ model of Bollerslev (1986) that is given by

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \tag{2.4}$$

where $\alpha_0 > 0$, $\alpha_1 \geq 0$, $\beta_1 \geq 0$, and $(\alpha_1 + \beta_1) < 1$; see Barone-Adesi et al. (1999), McNeil and Frey (2000), Nystrom and Skoglund (2002), Angelidis et al. (2005), Christoffersen and GonÇalves (2005), Giannopoulos and Tunaru (2005), Kuester et al. (2006) and Bali and Theodossiou (2007) among many others[2].

The basic $GARCH(1,1)$ model in (2.4) has been extended in several directions to cope with features of returns observed when analyzing real data. One of the most interesting of these features is the asymmetric response of volatility to positive and negative returns. The volatility is larger when past returns are negative than when they are positive; see Black (1976). This characteristic is known as leverage effect. Hentschel (1995) proposes the following specification of the conditional variance which nests several popular $GARCH$ specifications with leverage effect

$$\frac{\sigma_t^\delta - 1}{\delta} = \alpha_0 + \alpha_1 \sigma_{t-1}^\delta g^\nu\left(\epsilon_{t-1}\right) + \beta_1 \frac{\sigma_t^\delta - 1}{\delta} \tag{2.5}$$

where $g\left(\epsilon_t\right) = \left|\epsilon_t - b\right| - c\left(\epsilon_t - b\right)$. Among the most useful models encompassed by model (2.5) and usually implemented to estimate the $VaR$, one can find the Exponential $GARCH$ ($EGARCH$) of Nelson (1991) and the Asymmetric Power ARCH ($APARCH$) of Ding et al. (1993); for example, Angelidis et al. (2005) and Bali and Theodossiou (2007) who conclude that the $EGARCH$ model has the best performance when estimating the $VaR$ while Giot and Laurent (2003) fits the $APARCH$ model. The $EGARCH$ model is obtained from (2.5) when $\delta = 0$, $\nu = 1$ and $b = 0$. When

[2]The popular Riskmetrics model for the conditional variance is the $GARCH(1,1)$ model in (2.4) with the restriction $\alpha + \beta = 1$; see Longerstaey and More (1995) and Morgan (1995).

$\delta = \nu$, $b = 0$ and $|c| \leq 1$, we obtain the $APARCH$ model. There are another two very popular models that can be obtained as particular cases of the $APARCH$ model. When the power parameters are $\delta = \nu = 1$, the Threshold $GARCH$ ($TGARCH$) of Zakoian (1994) is obtained and when $\delta = \nu = 2$, we obtain the $GJR$ model of Glosten et al. (1993). Taking into account that the estimates of the power parameters are usually very close to 1, the results obtained from the $APARCH$ and $TGARCH$ models are usually very similar; see Rodríguez and Ruiz (2009).

The third component needed to compute the $VaR$ in equation (2.2) is $q_\alpha$ which is obtained from the distribution of the standardized returns $\epsilon_t$. There are two main alternatives to obtain the value of $q_\alpha$. First, it is possible to assume a particular distribution for $\epsilon_t$ and, consequently, $q_\alpha$ will be the $\alpha th$ quantile of this distribution. The most popular one is Normality; see Morgan (1995), Giot and Laurent (2003) and Bali and Theodossiou (2007) among many others. However, it has often been observed that when the conditional variance is specified as a $GARCH$-type model, the distribution of $\epsilon_t$ has fat tails. Therefore, when estimating the $VaR$, several authors have proposed leptokurtic distributions of $\epsilon_t$; see, for example, Pownall and Koedijk (1999), Mittnik and Paolella (2000), Manganelli and Engle (2001) and Angelidis et al. (2005). These authors generally assume that the distribution of $\epsilon_t$ is a standardized Student-$\nu$ or a GED distribution. Furthermore, to introduce skewness into the marginal distribution of returns several authors have proposed asymmetric conditional distributions of $\epsilon_t$[3]. For example, Giot and Laurent (2003) propose the standardized skewed-Student distribution of Hansen (1994) given by

$$f\left(\epsilon_t|\xi,\nu\right) = \begin{cases} \dfrac{2}{\xi + \dfrac{1}{\xi}} sg\left[\xi\left(s\epsilon + m\right)|\nu\right] & \text{if} \quad \epsilon < -\frac{m}{s} \\ \dfrac{2}{\xi + \dfrac{1}{\xi}} sg\left[\left(s\epsilon + m\right)/\xi|\nu\right] & \text{if} \quad \epsilon \geq -\frac{m}{s} \end{cases} \tag{2.6}$$

where $g(.|\nu)$ is the standardized Student density with $\nu$ degrees of freedom, $\xi$ is the coefficient of asymmetry, and $m$ and $s^2$ are the mean and the variance of the non-

[3]Alternatively, He et al. (2008) propose to introduce skewness in the marginal distribution of returns by assuming an asymmetric conditional mean.

standardized skewed Student given by $m = \frac{\Gamma\left(\frac{\nu-1}{2}\right)\sqrt{\nu-2}}{\sqrt{\pi}\Gamma\left(\frac{\nu}{2}\right)}\left(\xi - \frac{1}{\xi}\right)$ and $s^2 = \left(\xi^2 + \frac{1}{\xi^2} - 1\right) - m^2$, respectively. When $\xi > 0$, the density is skewed to the right while, when $\xi < 0$, it is skewed to the left. Lambert and Laurent (2000) show that the $100\alpha\%$ quantile of the standardized skewed-Student density is given by $q_\alpha = \frac{q_\alpha^* - m}{s}$, where $q_\alpha^*$ is the corresponding quantile of the skewed-Student density given by

$$q_\alpha^* = \begin{cases} \frac{1}{\xi} t_\alpha \left[\frac{\alpha}{2}\left(1+\xi^2\right)\right] & \text{if } \alpha < \frac{1}{1+\xi^2} \\ -\xi t_\alpha \left[\frac{1-\alpha}{2}\left(1+\xi^{-2}\right)\right] & \text{if } \alpha \geq \frac{1}{1+\xi^2} \end{cases}$$

and $t_\alpha$ is the $100\alpha\%$ quantile of the standardized Student-$\nu$ density. As an illustration, Figure 2.1 plots the skewed-Student distribution for different degrees of freedom and asymmetry parameter $\xi = 0.75, -0.75$. For small values of $\nu$ the density is more peaked and it becomes flatter as long as it increases.

Another asymmetric distribution is the skewed-generalized-t (SGT) distribution proposed by Theodossiou (1998). The SGT distribution has the attractive of encompassing most of the distributions usually assumed for standardized returns. For example, the Normal, GED, Student-$\nu$ and skewed-Student-$\nu$ distributions can be obtained as particular cases. However, in our experience, the maximization of the log-likelihood based on a SGT distribution is very complicated. Consequently, we will not consider further this distribution; see Bali and Theodossiou (2007) for an application of the SGT distribution in the estimation of the $VaR$.

Alternatively, instead of assuming a particular distribution of $\epsilon_t$, several authors propose to estimate directly $q_\alpha$. For example, Danielsson and de Vries (2000) and McNeil and Frey (2000) among others, use Extreme Value Theory ($EVT$) for the tails of the distribution of the standardized residuals; see Chan and Gray (2009) for a nice description of $EVT$ and its application to estimate the $VaR$ of daily electricity prices. This procedure is based on taking into account that, when the conditional mean and variance are correctly specified, the standardized residuals, $\widehat{\epsilon}_t = \frac{R_t - \widehat{\mu}_t}{\widehat{\sigma}_t}$ are $iid$. Then, they can be used to build the distribution function of the tail. If $F$ is the distribution of standardized returns, the excess distribution above the threshold $u$ is given by $F_u(y) = P[\widehat{\epsilon}_t - u \leq y \mid \widehat{\epsilon}_t > u] = \dfrac{F(y+u) - F(u)}{1 - F(u)}$. Therefore,

$$1 - F(\widehat{\epsilon}_t) = (1 - F(u))(1 - F_u(\widehat{\epsilon}_t - u)). \tag{2.7}$$

The function $(1 - F(u))$ can be estimated by the proportion of observations over the threshold, i.e. by $N/T$, where $N$ is the number of observations in the sample that exceed $u$. On the other hand, $1 - F_u(\widehat{\epsilon}_t - u)$ can be estimated by ML by assuming that the excess residuals over the threshold have, for example, a Generalized Pareto distribution ($GPD$) given by

$$G_{\xi,\beta}(y) = \begin{cases} 1 - (1 + \xi y/\beta)^{-1/\xi} & \text{if } \xi \neq 0 \\ 1 - \exp(-y/\beta) & \text{if } \xi = 0 \end{cases}$$

where $\beta$ is the scale parameter, $\xi$ is the shape parameter such that if $\xi > 0$ the distribution has heavy tails. The corresponding probability density function is given by

$$g_{\xi,\beta}(y) = \frac{1}{\beta}\left[1 + \frac{\xi y}{\beta}\right]^{-\dfrac{1+\xi}{\xi}}.$$

In practice, we fix the number of observations in the tail to be $N = k$, where $k << T$, obtaining a threshold at the $(k+1)\,th$ order statistic. Consequently, if $\widehat{\epsilon}_{(1)} \geq ... \geq \widehat{\epsilon}_{(T)}$ are the ordered standardized residuals, the threshold is $\widehat{\epsilon}_{(k+1)}$ and the GPD is fitted to $\left(\widehat{\epsilon}_{(1)} - \widehat{\epsilon}_{(k+1)}, ..., \widehat{\epsilon}_{(k)} - \widehat{\epsilon}_{(k+1)}\right)$. Using (2.7) we get the following tail estimator for $\widehat{\epsilon}_t > \widehat{\epsilon}_{(k+1)}$

$$\widehat{F}(\widehat{\epsilon}_t) = 1 - \frac{k}{T}\left(1 + \widehat{\xi}\frac{\widehat{\epsilon}_t - \widehat{\epsilon}_{(k+1)}}{\widehat{\beta}}\right)^{-1/\widehat{\xi}}. \tag{2.8}$$

Finally, if $\alpha < \dfrac{k}{T}$, the quantile can be obtained from (2.8) as follows

$$\widehat{q}_\alpha = -\left(\widehat{\epsilon}_{(k+1)} + \frac{\widehat{\beta}}{\widehat{\xi}}\left(\left(\frac{\alpha}{k/T}\right)^{-\widehat{\xi}} - 1\right)\right). \tag{2.9}$$

The use of the $GPD$ for the excess residuals is just an example of a heavy-tailed distribution. Gnedenko (1943) characterizes all such distributions by the following formula for $x > u$

$$1 - F(x) = x^{-1/\xi}L(x). \tag{2.10}$$

Applying expression (2.10) to the ordered residuals beyond the $(k+1)\,th$ order statistic and choosing $L(\widehat{\epsilon}) = \dfrac{k}{T}\left(\widehat{\epsilon}_{(k+1)}\right)^{1/\xi}$, the following tail distribution is obtained

$$F\left(\widehat{\epsilon}_t\right) = 1 - \frac{k}{T}\left(\frac{\widehat{\epsilon}_t}{\widehat{\epsilon}_{(k+1)}}\right)^{-1/\xi}. \tag{2.11}$$

In this case, the shape parameter $\xi$ can be estimated using the estimator proposed by Hill (1975) that is given by

$$\widehat{\xi}^{(H)} = \frac{1}{k}\sum_{j=1}^{k} \log\left(\widehat{\epsilon}_{(k)}\right) - \log\left(\widehat{\epsilon}_{(k+1)}\right). \tag{2.12}$$

The estimation of the quantile is then

$$\widehat{q}_\alpha = -\widehat{\epsilon}_{(k+1)}\left(\frac{\alpha}{k/T}\right)^{-\widehat{\xi}^{(H)}}. \tag{2.13}$$

One important issue related with the Hill estimator is the choice of the number of observations in the tail, $k$. In this sense, McNeil and Frey (2000) show that the $EVT$ method based on the $GPD$ distribution gives more stable quantile estimates than the Hill estimator. To illustrate this point, the quantile is obtained by the Hill estimator and the $GPD$ distribution using the procedure explained above. Figure 2.2 plots Hill estimates of the 1% quantile of the returns of the S&P500 index observed from 29/08/1995 to 20/10/2005 for different values of $k$. This figure shows that, as expected, the Hill estimator of $q_{0.99}$ is an increasing function of $k$, although it is very unstable when the number of observations over the threshold is small. This figure also plots the estimates based on the $GPD$ distribution which are also very unstable for small $k$. However, when $k > 30$, the estimate of $q_{0.99}$ is approximately 2.6 regardless of $k$. Note that the same estimate is obtained by the Hill estimator when $30 < k < 250$. Only for very large values of $k$ the Hill estimator generates estimates of $q_{0.99}$ well over 2.6. Therefore, if the number of observations in the tail is moderate, i.e. between 30 and 250, both estimators should give the same answer. It is clear that the choice of the threshold is an important issue in $EVT$ since it may have severe consequences on the tail estimates. Danielsson and de Vries (1997) and Danielsson et al. (2001) have developed bootstrap methods for optimal threshold selection in the context of the Hill estimator. However, the selection of the threshold using bootstrap procedures is very time consuming. Alternatively, Gonzalo and Olmo (2004) propose a single step approach to threshold selection.

Chan et al. (2007) derive the asymptotic distribution of the quantile estimator of McNeil and Frey (2000) in (2.9) without assuming a specific parametric distributional assumption on the heavy tailed distribution of $\epsilon_t$. Then, they propose two alternative methods to construct confidence intervals of the $VaR$. The first method is the traditional method based on the asymptotic Normality of the $VaR$ estimator. Alternatively, they propose to construct confidence intervals by the tilting method of Hall and Yao (2003) and Peng and Qi (2003). Note, that the confidence intervals for the $VaR$ constructed in this way do not incorporate the uncertainty due to the estimation of the parameters of the conditional mean and standard deviation.

Alternatively, the quantile $q_\alpha$ can be estimated using bootstrap methods that do not assume any particular distribution of the errors and incorporate the uncertainty of the estimated parameters; see Ruiz and Pascual (2002b) for a review of the literature on using bootstrap procedures in financial time series and, in particular, for the estimation of the $VaR$. In particular, Hull and White (1998) and Barone-Adesi et al. (1999) propose a bootstrap method called Filtered Historical Simulation ($FHS$) based on using random draws with replacement from the standardized residuals. Bootstrap procedures have the advantage of allowing to obtain confidence intervals for the estimated $VaR$. For example, Christoffersen and GonÇalves (2005) implement the bootstrap procedures proposed by Pascual et al. (2006) to obtain confidence intervals for the $VaR$ that incorporate the parameter uncertainty. They show that the confidence intervals for $HS$ are too narrow and do not contain the true $VaR$ with the desire frequency while the methods that properly account for conditional variance dynamics imply confidence intervals with coverages close to the nominal. Bootstrap procedures have also been implemented by Hartz and Paolella (2006) who additionally propose a bias-correction method for improving the $VaR$ forecasting ability of the $GARCH$ model with $Normal$ errors ($GARCH-N$).

The procedures described up to now are based on assuming a parametric specification of the conditional mean and variance. However, semiparametric and nonparametric specifications have also been considered in the literature. For example, Fan and Gu (2003) introduce a semiparametric model to estimate the volatility using the geometric Brownian motion, a time-dependent diffusion model, as a discretization of the $IGARCH(1,1)$ model of Riskmetrics. In order to estimate the decay factor needed for the Riskmetrics methodology they propose two alternatives, one resulting in a data dependent decay factor which remains constant in the forecasting period, and the other which adapts automatically to changes in stock price dynamics, adding flexibility to the first decay factor. Additionally, Fan and Gu (2003) propose a symmetric nonparametric estimation approach to estimate the quantiles of the standardized residuals. On the other hand, Martins-Filho and Yao (2006), based on the two stage approach of McNeil and Frey (2000), propose a nonparametric estimation procedure for the conditional mean and variance using the local linear estimator of Fan (1992). Furthermore, they propose a method based on L-Moment theory instead of the $GPD$ used by McNeil and Frey (2000). These nonparametric methods are more difficult to implement than the parametric procedures. However, inferential gains can be obtained when the assumptions of the parametric models are wrong. Another nonparametric procedure is that developed by Chen and Tang (2005b). They propose to calculate the $VaR$ by implementing kernel smoothing on the empirical distribution of returns in such a way that the estimator of the $VaR$ is a weighted average of the order statistics around $R_{\omega:T}$. They also emphasize the importance of the standard error of the $VaR$ estimates and develop a procedure for its estimation based on a kernel estimation of the spectral density function of a series built using the smoother function. More recently, Cai and Wang (2008) developed a nonparametric estimator of the $VaR$ and the $ES$ by obtaining a weighted double kernel local linear estimator of the conditional distribution function. The proposed estimator is a combination of the weighted Nadaraya-Watson method of Cai (2002) and the double kernel local linear method of Yu and Jones (1998).

Finally, there is another alternative to estimate the $VaR$ by modeling directly the dynamic evolution of the quantiles over time. The Conditional Autoregressive Value at Risk ($CAViaR$) was introduced by Engle and Manganelli (2004) who propose the following equation for the $VaR$

$$VaR_t^\alpha = \beta_0 + \beta_1 VaR_{t-1}^\alpha + l\left(\beta_2, R_{t-1}, VaR_{t-1}^\alpha\right) \tag{2.14}$$

where different forms of the function $l$ can be proposed. Some examples are the asymmetric slope, $l(\cdot) = \beta_2 (R_{t-1})^+ + \beta_3 (R_{t-1})^-$, where $(x)^+ = \max(x, 0)$, $(x)^- = -\min(x, 0)$, and the adaptive, $l(\cdot) = \beta_2 \left\{ \left[1 + \exp\left(G\left[R_{t-1} + VaR_{t-1}^\alpha\right]\right)\right]^{-1} - \alpha \right\}$, where G is some positive finite number. The parameters of this model are estimated by the method of regression quantiles developed by Koenker and Bassett (1978). Manganelli and Engle (2001) also incorporate $EVT$ to $CAViaR$. The procedure is the following: first, fit a $CAViaR$ model to get an estimation of the $VaR$ for a large $\alpha$, for example 10%, then construct the series of standardized quantile residuals as follows: $\dfrac{\widehat{\epsilon}_{t,\alpha}}{\widehat{VaR}_t^\alpha} = \left(\dfrac{R_t}{\widehat{VaR}_t^\alpha}\right) - 1$ and apply $EVT$ to this series to get an estimation of the tail $\widehat{q}_p$ for $p < \alpha$. Then, the $VaR$ is calculated as

$$\widehat{VaR}_t^p = \widehat{VaR}_t^\alpha \left(1 + \widehat{q}_p\right). \tag{2.15}$$

Alternatively, DeRossi and Harvey (2006) propose to combine the approach of Engle and Manganelli (2004) with signal extraction. The idea is to use some of the forms of the function $l$ and approximate them to the filtered estimators of time-varying quantiles. Recently, Gourieroux and Jasiak (2006) proposed a dynamic adaptive quantile which improves the approach of Engle and Manganelli (2004) by taking into account the monotonicity of quantile estimators which ensures that the quantile is an increasing function of $\alpha$. On the other hand, Chen and Chen (2005) compare the performance of the Riskmetrics approach and the $GARCH(1,1) - N$ and $GARCH(1,1) - t$ models for estimating $VaR$ with the combination of them with quantile regression. They conclude that the quantile regression combined with the $GARCH(1,1) - t$ provides the best estimates.

### 2.2.2 Backtesting VaR estimates

In order to asses the accuracy of $VaR$ estimates, the Basel Committee on Banking Supervision (1996b) and the amendments of Basel Committee on Banking Supervision (1996a) develop a statistical testing device denominated backtesting. According to their requirements, the backtesting should be based on 250 one step-ahead estimates of the $VaR$, i.e. estimates over one year. In this section, we review the most popular backtesting procedures proposed in the literature. Backtesting is based on testing whether the $VaR$ estimates are statistically accurate. When there are several alternative estimators of the $VaR$, one may additionally want to choose the best among those that generate accurate estimates; see, for example, Sarma et al. (2003) and Angelidis and Degiannakis (2007).

Backtesting procedures are based on the failure process $I_t^\alpha = 1\left(R_t < -VaR_t^\alpha\right)$, $t = T+1, ...T+n$, where $1\left(\cdot\right)$ is the indicator function, $T$ is the size of the sample used to estimate the $VaR$ and $n$ is the number of one step-ahead $VaR$'s computed. A $VaR$ estimator is accurate if and only if

$$\underset{t-1}{E}\left[I_t^\alpha\right] = \alpha. \tag{2.16}$$

Most backtesting procedures are based on testing some of the implications of this condition. The most popular backtesting procedure, proposed by Kupiec (1995), is based on the number of failures defined as $x = \sum\limits_{T+1}^{T+n} I_t^\alpha$ which has a binomial distribution with parameters $n$ and $\alpha$. Kupiec (1995) proposes to test the null hypothesis $H_0$ : $E\left[I_t^\alpha\right] = \alpha$, using the following likelihood ratio statistic

$$LR_{uc} = 2\log\left[\left(1-\frac{x}{n}\right)^{n-x}\left(\frac{x}{n}\right)^{x}\right] - 2\log\left[\left(1-\alpha\right)^{n-x}\alpha^{x}\right]. \tag{2.17}$$

Under the null, the $LR_{uc}$ statistic has asymptotically a $\chi^2_{(1)}$ distribution. It has low power when implemented with small samples. However, note that the null hypothesis is testing whether the unconditional expectation is $\alpha$, which is not the hypothesis of interest in (2.16). Consequently, Christoffersen (1998) proposes a test of conditional coverage, where the null hypothesis is given by $H_0 : E\left[I_t^\alpha | I_{t-1}^\alpha\right] = \alpha$. This is equivalent to testing whether $I_t^\alpha$ are *iid Ber* $(\alpha)$ random variables against the alternative of first order Markov dependence. Note that this condition is necessary but not sufficient for the hypothesis in (2.16). This test considers whether the unconditional coverage is correct and adds a term to consider the serial independence of the failure process $\{I_t\}$. The serial independence term, $LR_{ind}$, is defined as follows

$$LR_{ind} = 2\log\left[(1-\pi_{01})^{n_{00}}\pi_{01}^{n_{01}}(1-\pi_{11})^{n_{10}}\pi_{11}^{n_{11}}\right] - 2\log\left[(1-\pi)^{n_{00}+n_{10}}\pi^{n_{01}n_{11}}\right] \quad (2.18)$$

where $n_{ij}$ is the number of $I_t^\alpha$ observations with value $i$ followed by an observation with value $j$, for $i,j = 0,1$ and $\pi_{01} = \frac{n_{01}}{n_{00}+n_{01}}$, $\pi_{11} = \frac{n_{11}}{n_{10}+n_{11}}$. Under the null hypothesis $\pi_{01} = \pi_{11} = \pi = \frac{n_{01}+n_{11}}{n}$ and the $LR_{ind}$ statistic has an asymptotic $\chi^2_{(1)}$ distribution. Finally, the likelihood ratio for conditional coverage, $LR_{cc}$ is defined as

$$LR_{cc} = LR_{uc} + LR_{ind} \quad (2.19)$$

which has asymptotically a $\chi^2_{(2)}$ distribution under the null.

Recently, other tests for independence based on the autocovariances $Cov\left(I_t^\alpha, I_{t-j}^\alpha\right)$ have been proposed. For example, Berkowitz et al. (2006) discuss the following Portmanteau test

$$LB(m) = (n)(n+2)\sum_{j=1}^{m}(n-j)^{-1}r_j^2 \quad (2.20)$$

where $r_j$ is the order $j$ sample autocorrelation of $I_t^\alpha - \alpha$. Under the null, $LB(m)$ is asymptotically $\chi^2_{(m)}$. On the other hand, Engle and Manganelli (2004) suggest a dynamic quantile ($DQ$) test obtained by regressing $I_t^\alpha - \alpha$ against its lagged variables and other values included in the conditioning set and testing whether these variables are significant.

The backtesting tests described above are based on the assumption that the parameters of the models fitted to estimate the $VaR$ are known. However, in practice, these parameters have to be estimated. Escanciano and Olmo (2010) show that the use of standard unconditional and independence backtesting procedures to asses $VaR$ models in out-of-sample environments can be misleading. They quantify the risk associated with the estimation of the parameters in a very general class of dynamic parametric

$VaR$ models and propose a correction of the standard backtesting procedures that takes into account such a risk. They show that one of the main determinants of the corrected asymptotic variance is the forecasting scheme used to generate the forecasts of the $VaR$, i.e. whether one uses recursive, rolling or fix parameter estimates.

The backtesting procedures help to decide whether a particular procedure gives accurate estimates of the $VaR$. However, when several accurate estimators are available, one wants also to decide which estimator is best among them. With this goal, Lopez (1999) proposes to choose the procedure that minimizes $C_m = \sum_{t=T+1}^{T+n} C_{m,t}$ where

$$C_{m,t} = \begin{cases} f\left(R_t, VaR^{\alpha}_{m,t}\right) & \text{if } R_t < VaR^{\alpha}_{m,t}, \\ g\left(R_t, VaR^{\alpha}_{m,t}\right) & \text{if } R_t \geq VaR^{\alpha}_{m,t}. \end{cases}$$

where the index $m$ is used to represent the procedure $m$ to estimate the $VaR$ and $f(x, y)$ and $g(x, y)$ are functions such that $f(x, y) \geq g(x, y)$.

Different loss functions has been proposed in the literature; see Lopez (1999). Sarma et al. (2003) and Angelidis and Degiannakis (2007) use the following Regulatory Loss function ($RLF$) that is similar to the Quadratic Loss Function proposed by Lopez (1999)

$$C_{m,t} = \begin{cases} \left(R_t - VaR^{\alpha}_{m,t}\right)^2 & \text{if } R_t < VaR^{\alpha}_{m,t}, \\ 0 & \text{if } R_t \geq VaR^{\alpha}_{m,t}. \end{cases} \tag{2.21}$$

Angelidis et al. (2005) proposed the Quantile Loss function ($QLF$) that additionally penalizes for higher than needed amount of capital. It is defined by

$$C_{m,t} = \begin{cases} \left(R_t - VaR^{\alpha}_{m,t}\right)^2 & \text{if } R_t < VaR^{\alpha}_{m,t}, \\ \left(R_{\omega:n} - VaR^{\alpha}_{m,t}\right)^2 & \text{if } R_t \geq VaR^{\alpha}_{m,t}. \end{cases} \tag{2.22}$$

On the other hand, Giacomini and Komunjer (2005) and Bao et al. (2006) compare competing $VaR$ forecasts using the predictive quantile loss function ($PQLF$) based on the methodology of Koenker and Bassett (1978). The $PQLF$ function is given by

$$C_{m,t} = \left[\alpha - 1\left(R_t < VaR^{\alpha}_t\right)\right]\left[R_t - VaR^{\alpha}_t\right]. \tag{2.23}$$

Alternatively, when trying to establish the superiority between two models Sarma et al. (2003) propose the following testing procedure

$$H_0 : \{\theta = 0\} \ vs \ H_1 : \{\theta < 0\}$$

where $\theta$ is the median of the distribution of the loss differential between procedure $i$ and procedure $j$, $z_t = C_{i,t} - C_{j,t}$. The number of non-negative $z's$ is defined as $S_{ij} = \sum_{t=T+1}^{T+n} \psi_t$, where $\psi_t = 1\,(z_t \geq 0)$. Under the null hypothesis, the exact distribution of $S_{ij}$ is binomial with parameters $(n, 0.5)$ and the asymptotic distribution of the standardized $S_{ij}$ is given by

$$\frac{S_{ij} - 0.5n}{\sqrt{0.25n}} \overset{a}{\sim} N\,(0,1)\,; \tag{2.24}$$

see Diebold and Mariano (1995). If $H_0$ is rejected, model $i$ is significantly better than model $j$ for the chosen loss function. Note that the statistic in (2.24) can be obtained as the t-statistic of the regression of $z_t$ on a constant using the Newey and West (1987) heteroscedasticity autocorrelation consistent (HAC) standard errors.

An alternative to the test in (2.24) is the test of conditional predictive ability, proposed by Giacomini and White (2006), which takes into account the estimation uncertainty due to model selection. The one-step-ahead conditional predictive ability statistic is given by

$$CPA_{T,n} = n\left(n^{-1}\sum_{t=T}^{T+n-1} h_t \Delta C_{m,t+1}\right)' \widehat{\Omega}_n^{-1}\left(n^{-1}\sum_{t=T}^{T+n-1} h_t \Delta C_{m,t+1}\right), \tag{2.25}$$

where $\widehat{\Omega}_n$ is a consistent estimate of the variance of $h_t \Delta C_{m,t+1}$ and $h_t$ is a $q \times 1$ vector given by $h_t = (1, \Delta C_{m,t})'$. The null hypothesis of equal conditional predictive ability is rejected when $CPA_{T,n} > \chi^2_q$.

On the other hand, Angelidis and Degiannakis (2007) and Bao et al. (2006) propose to compare alternative models using the test of superior predictive ability ($SPA$) of Hansen (2005). The null hypothesis, that the benchmark model ($m = 0$) is not inferior than the alternatives, is tested with the following statistic

$$T_n^{SPA} = \max\left[\max_{m=1,\ldots,M} \frac{n^{1/2}\overline{z}_m}{\widehat{\omega}_m}, 0\right], \tag{2.26}$$

where $\widehat{\omega}_m^2$ is a consistent estimator of $\omega_m^2 = var\left(n^{1/2}\overline{z}_m\right)$, $\overline{z}_m = n^{-1}\sum_{t=1}^{n} z_{m,t}$ and $z_{m,t} = C_{0,t} - C_{m,t}$. The estimation of $\omega_m^2$ and the p-values of the $T_n^{SPA}$ can be obtained using the stationary bootstrap of Politis and Romano (1994) with the optimal block-size chosen by the block selection algorithm proposed by Politis and White (2004).

Finally, Giacomini and Komunjer (2005) propose to compare two $VaR$ forecasts versus its combination using a conditional quantile forecast encompassing test of superior predictive ability. A rejection of the test provides statistical evidence that the combination outperforms the two individual forecasts.

## 2.3 Estimation and testing of Expected Shortfall ($ES$)

This section describes different methods proposed in the literature for estimating $ES$. As we mentioned above, $ES$ is a relatively new measure of risk, and consequently there are fewer papers dealing with its estimation. Most of the papers actually estimate $CVaR$ instead of $ES$. Remember that the $CVaR$ only is coherent if the returns have a continuous probability distribution but, in practice, the distribution of returns is often assumed to be continuous and, in this case, the $CVaR$ and the $ES$ are equivalent. We also describe methods to evaluate the accuracy of the estimated $ES$.

### 2.3.1 Estimation

Acerbi and Tasche (2002) propose to estimate the $ES$ using the $VaR$ estimator based on Historical Simulation. In this case, the estimator is given by

$$\widehat{ES}_t^{\alpha} = \overline{R}_{(\omega)}, \tag{2.27}$$

where $\overline{R}_{(\omega)} = \dfrac{\sum_{i=1}^{\omega} R_{i:T}}{\omega}$ is the average of the smallest $100\alpha\%$ returns. This estimator has a positive bias attributable to the negative biases of the order statistics. Consequently, Inui and Kijima (2005) has proposed an extrapolation method to adjust the bias and stabilize the estimator.

Alternatively, several authors propose to estimate the $ES$ as the average of observed returns beyond the $VaR$ when the $VaR$ has been estimated by one of the methods described in the previous section; see, for example, Giot and Laurent (2003) and Bali and Theodossiou (2007).

On the other hand, instead of using the sample mean of the returns beyond the $VaR$, note that if returns are given by equation (2.1) then the $ES$ is given by

$$ES_t^{\alpha} = \mu_t + \sigma_t \underset{t-1}{E} \left[\epsilon_t | \epsilon_t \leq q_\alpha\right]. \tag{2.28}$$

There are different alternative methods proposed to calculate $\underset{t-1}{E} \left[\epsilon_t | \epsilon_t \leq q_\alpha\right]$. First, one can assume a particular distribution for the innovations and calculate analytically the corresponding expectation. If, for example, they are Normal, then $E\left[\epsilon_t | \epsilon_t \leq q_\alpha\right] = -\frac{\phi\left(\Phi_\alpha^{-1}\right)}{\alpha}$, where $\Phi_\alpha^{-1}$ is the $\alpha th$ quantile of the standard Normal distribution. On the other hand, if the innovations are Student-$\nu$, then

$$E\left[\epsilon_t | \epsilon_t \leq q_\alpha\right] = \frac{\nu - 2}{\alpha\left(1 - \nu\right)} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\left(\nu - 2\right)\Gamma\left(\frac{\nu}{2}\right)}} \left(1 + \frac{q_\alpha^2}{\nu - 2}\right)^{\frac{1-\nu}{2}}$$

where $q_\alpha$ is the $\alpha th$ quantile of the standardized Student-$\nu$; see Christoffersen and GonÇalves (2005) for an empirical application. In the case of the $GED$ and the Skewed-t distributions, the corresponding conditional expectations can be estimated by Monte Carlo simulations.

Another procedure to estimate the $ES$ is by using $EVT$ in order to estimate the tail of the distribution of the standardized residuals and then, calculate the conditional expectation of the values beyond the quantile $q_\alpha$; see McNeil and Frey (2000). In this case, if the excess residuals over the threshold $u$ are assumed to follow a $GPD$ distribution with parameters $\xi < 1$ and $\beta$, then the expected shortfall is estimated as follows

$$\widehat{ES}_{t+1}^{\alpha} = \widehat{\mu}_{t+1} + \widehat{\sigma}_{t+1}\widehat{q}_\alpha \left(\frac{1}{1-\widehat{\xi}} + \frac{\widehat{\beta} - \widehat{\xi}\epsilon_{(k+1)}}{\left(1-\widehat{\xi}\right)\widehat{q}_\alpha}\right). \tag{2.29}$$

Alternatively, using the Hill estimator we can obtain the following estimation of the $ES$

$$\widehat{ES}_{t+1}^{\alpha} = \widehat{\mu}_{t+1} + \widehat{\sigma}_{t+1}\frac{\widehat{q}_\alpha}{1-\widehat{\xi}^{(H)}}. \tag{2.30}$$

Another method for estimating $ES$ is $FHS$ as in Giannopoulos and Tunaru (2005). Once the bootstrap distribution of returns is obtained, the $ES$ estimator is calculated as the sample average of the returns that exceed $VaR$. We can also obtain the bootstrap distribution of returns using the procedure develop by Pascual et al. (2006) and then estimate the $ES$ as the average of the returns that exceed $VaR$. Christoffersen and GonÇalves (2005) implement the bootstrap procedure of Pascual et al. (2006) to obtain confidence intervals for alternative estimators of the $ES$. They show that $ES$ measures are generally less accurate than $VaR$ measures and that the confidence bands around $ES$ are also less reliable. Table 2.2 contains a summary of different assumptions on $\mu_t$, $\sigma_t$ and the distribution of $\epsilon_t$ proposed in the literature for estimating the $ES$.

### 2.3.2 Backtesting ES estimates

In order to evaluate the adequacy of the estimated $ES$, Angelidis and Degiannakis (2007) propose a two steps evaluation framework that extends the evaluation approach of Lopez (1999). In the first step, the correct conditional coverage of the $VaR$ forecast is tested using the $LR_{cc}$ statistic in equation (2.19). In the second step, the loss function is calculated with respect to the $ES$ instead of the $VaR$, because the $VaR$ does not give information about the size of the expected loss. The loss function is then

$$C_{m,t} = \begin{cases} \left(R_t - ES^{\alpha}_{m,t}\right)^2 & \text{if } R_t < VaR^{\alpha}_t \\ 0 & \text{if } R_t \geq VaR^{\alpha}_t \end{cases}, \tag{2.31}$$

where the subindex $m$ refers to model $m$. For each model, the mean squared error is calculated by $MSE = \dfrac{1}{n} \sum_{t=T+1}^{T+n} C_{m,t}$. We choose the model that minimizes the $MSE$.

However, in this case, there is some ambiguity about the interpretation of the $MSE$. In order to overcome this problem, alternative models can be tested by using the test of superior predictive ability of Hansen (2005). Note that this procedure tests for superiority among the models that provide accurate $VaR$ forecasts, but it is not a method to test the accuracy of the $ES$ forecasts.

## 2.4 Empirical Application: Estimating the $VaR$ and $ES$ of S&P500 returns.

### 2.4.1 Model fitting

In this section, we implement the methods described above to estimate the $VaR$ and $ES$ of a series of daily returns of the S&P500 index observed from 29/08/1995 to 20/10/2005. The series of returns, plotted in the first panel of Figure 2.3, show volatility clustering. Table 2.3, that reports some descriptive statistics, shows that the returns have excess of kurtosis and skewness. This table also reports the ratio between the Box-Pierce statistic and its corresponding asymptotic 5% critical value for testing whether the first 20 autocorrelations are jointly equal to zero. We can observe that the ratio is smaller than one and consequently, the null is not rejected. The second panel of Figure 2.3 plots the correlogram of the series of returns together with their 95% confidence bands computed as suggested by Diebold (1988) to account for the presence of conditional heteroscedasticity. The dependence in $S\&P500$ returns seems to be well represented by a white noise. Table 2.3 also reports the ratio with respect to its asymptotic critical value of the Rodriguez and Ruiz (2005) statistic for testing whether the first 20 autocorrelations of absolute returns are jointly equal to zero[4]. In this case, the null is clearly rejected. This result is in concordance with the correlogram of absolute returns plotted in the third panel of Figure 2.3. The sample correlations of absolute returns are positive and highly persistent, being significantly different from zero even for very long lags. Therefore, S&P500 returns could be conditionally heteroscedastic. Figure 2.3 also plots the cross-correlogram between returns and squared returns, $Corr\left(y_t, y_{t+h}^2\right)$, $h = ..., -2, -1, 0, 1, 2...$ These cross-correlations suggest that the volatility of S&P500 seems to have a leverage effect. The evidence of an effect of the volatility in the conditional mean of returns is much weaker given that the cross-

[4]Rodriguez and Ruiz (2005) propose to test for conditional homoscedasticity by using the following statistic $Q_i^*(M) = T\sum_{k=1}^{M-i}\left[\sum_{l=0}^{i}\widetilde{r}(k+l)\right]^2$ where $\widetilde{r}(k+l)$ is the standardized sample autocorrelation of order $k+l$ of absolute returns, $M = [M/3] - 1$ is the number of autocorrelations and $T$ is the sample size. This statistics is more powerful than the more popular McLeod and Li (1983) test because it takes into account that under the null the sample autocorrelations have to be equal to zero and mutually uncorrelated.

correlations are not significant when $h < 0$.

Summarizing, the $S\&P500$ returns seem to be conditionally heteroscedastic with the volatility being larger when past returns are negative than when they are positive. Consequently, we fit several $GARCH$ type models with leverage effect, each of them with alternative assumptions on the error distribution. In particular, we consider the $GARCH$, $TGARCH$, $GJR$, $EGARCH$ and $APARCH$ models. The distributions of the errors assumed are $Normal$, $Student - \nu$, $GED$ and $Skewed - t$. Table 2.4 reports the maximum likelihood estimates of the parameters[5]. The asymmetry parameter of the $Skewed - t$ distribution is significant in all the models considered and its estimated value is always around $-0.11$. However, the estimated degrees of freedom of the Student-$\nu$ and of the Asymmetric Student-$\nu$ distributions are always larger than 10. Therefore, it seems that it is more important the asymmetry than the excess kurtosis of the errors. Figure 2.4 plots the estimated kernel densities of the standardized residuals when the $GARCH$ and $APARCH$ models are fitted together with each of the densities estimated. In both cases, it seems that the $GED$ and $Skewed - t$ distributions give the best fit. The results for all other models considered are very similar and not plotted to save space. Finally, the estimate of the power parameter in the $APARCH$ model is rather close to one.

Table 2.5 reports diagnostics on all the estimated models. In particular we report the skewness, kurtosis, ratio of the $Q^*_{[20/3]-1}(20)$ and $Q(20)$ statistics and the correlation of order one between standardized residuals and future squared standardized residuals. All the models are successful in explaining the autocorrelations of absolute returns which are not any longer significant. On the other hand, the cross-correlations are still significant. Additionally, when the distribution of the errors is assumed to be symmetric, the skewness of the standardized residuals is still different from zero. Finally, Table 2.5 shows that the kurtosis of the standardized returns is clearly smaller than that of the original returns reported in Table 2.3. However, it is still larger than the kurtosis of the distribution estimated. In any case, for any of the four distributions considered,

[5]The estimations were obtained by Matlab codes.

the kurtosis of the standardized returns is smaller when the $TGARCH$, $EGARCH$ and $APARCH$ models are fitted. For these three models, the kurtosis are very similar among them. Any of those models with $Skewed - t$ errors seem to provide the best concordance between the moments implied by the model and the empirical moments of the original returns.

### 2.4.2 Estimates of the $VaR$

For each of the models estimated, the $VaR^{\alpha}_{T+1}$, for $\alpha = 0.01$, has been computed by assuming that the conditional mean is zero and the conditional variance and error distribution are those implied by the estimated models reported in Table 2.4. The $VaR$ has also been estimated by assuming that the conditional variance is that estimated by assuming Normal errors and then, the distribution of the errors estimated by bootstrapping and by the Hill and $GPD$ procedures. For the $EVT$ method of McNeil and Frey (2000), we compute the maximum likelihood estimates of the $GPD$ distribution which are given by $\beta = 0.51$ and $\xi = 0.17$. Finally, we also estimate the $VaR$ using $HS$ and the asymmetric and adaptive versions of the $CAViaR$ model.

Figure 2.5 represents scatter-plots of the $VaR's$ estimated assuming $Normal$ errors with the five models considered. We can observe that the estimated $VaR$ obtained with the $TGARCH$, $EGARCH$ and $APARCH$ models are almost identical. Furthermore, the estimates of $VaR$ obtained by the $GJR$ model are also very similar. The only model that generates estimates clearly different from all others is the $GARCH$. Therefore, it seems that for a given error distribution the $TGARCH$, $EGARCH$ and $APARCH$ models generate similar estimates of the $VaR's$. A similar conclusion is obtained from Figure 2.6 that plots scatter-plots of the $VaR$ when the distribution of the errors is approximated by using bootstrap resampling. Therefore, it seems that using bootstrap procedures to approximate the distribution of the errors when estimating the quantile $q_\alpha$ does not account for differences in the specification of the conditional variances.

In Figure 2.7 we also represent scatterplots of the $VaR's$ estimated for the $EGARCH$ model with different assumptions on the error distribution. First of all, comparing Figures 2.5, 2.6 and 2.7, we can observe that the differences among estimated $VaR's$ are larger when the distribution is fixed and the model changes than when the model is fixed and the distribution changes. Therefore, it seems that it is more important to choose correctly the model for the conditional variance and that the particular distribution of the errors chosen only has a marginal effect on the estimates of the $VaR$. Figure 2.7 also shows that computing the $VaR$ using the $Hill$ or the $GPD$ extreme value estimators provide nearly identical results.

The analysis in the previous figures does not include estimates of the $VaR$ based on estimating directly the quantiles. Consequently, Figure 2.8 contains scatter-plots of the estimates of the Asymmetric $CAViaR$ against the $EGARCH$ estimates with all distributions of the errors respectively. This Figure shows that the $CAViaR$ is in general less conservative than the others, because, in most of the cases, the $VaR$ estimated using the $CAViaR$ Asymmetric is greater in absolute value than the $VaR$ estimated using the other models.

Finally, to have a clearer picture of the shape of the $VaR$ estimates and to include also in the analysis the estimates obtained using $HS$ and the Adaptive $CAViaR$, Figure 2.9 plots both $CAViaR$ estimates, $HS$ and the $EGARCH$ estimates with all the considered distributions. First of all, one can observe that $HS$ and Adaptive $CAViaR$ estimates are almost constant over time when compared with the estimates obtained by any of the alternative procedures. Furthermore, one can also observe that the Asymmetric $CAViaR$ has a shape similar to this observed in the $EGARCH$ estimates although the estimates are smaller most of the time, (confirming the conclusion from Figure 2.8). Finally, the rest of the models show almost the same behavior.

We backtest the models using the procedures described in the previous section with $n = 1000$. In particular, Table 2.6 reports the p-values of the backtesting tests of Christoffersen (1998) in (2.19), the $DQ$ test of Engle and Manganelli (2004) and the $LB$ test of Berkowitz et al. (2006) in (2.20) computed for $m = 5$ and 20. Only the Historical Simulation $VaR$ estimates are clearly rejected by all the tests. With respect to the rest of the $VaR$ estimates most of them are not rejected as inaccurate when using the $DQ$ and $LB$ tests. It seems that it is more useful to test for the accuracy of the $VaR$ estimates by using the Christoffersen (1998) test that is able to reject more models than the alternatives.

Next, we implement the tests described above to choose among those models that provide accurate $VaR$ estimates according to the test of Christoffersen (1998) when the significance level is 10%. First, we implement the procedure proposed by Lopez (1999) with the $RLF$, the $QLF$ and the $PQLF$ loss functions defined in (2.21), (2.22) and (2.23) respectively. Table 2.7 that reports the corresponding values of the functions for each of the models and for the different confidence levels, show that the model chosen as the best changes depending on the particular loss function chosen. Using the $RLF$ and the $QLF$, the model chosen is the $EGARCH$ with $Skewed-t$ errors, while according to the $PQLF$ the best model is the $GARCH$ with $Normal$ errors. However, it is important to remark that the values of the objective function under the $QLF$ and $PQLF$ are very similar and consequently, not very useful to distinguish among alternative estimators of the $VaR$; see Lopez (1999). Alternatively, Table 2.7 also reports the p-values of the $SPA$ test in (2.26) using the three loss functions considered in this Chapter. Every model is considered as the benchmark and compared with the others. We can observe that the $RLF$ is not very informative when using the $SPA$ test as none of the models is rejected as inferior. On the other hand, the models rejected as inferior when using the $QLF$ or the $PQLF$ functions are different. The models which are not rejected when using any of the three functions are the $GARCH-N$, $TGARCH-N$, $GJR-N$, $GJR-GED$, $EGARCH-Skewed$ and $APARCH-N$.

Table 2.8 reports the results of the test proposed by Sarma et al. (2003) to compare the models two by two with the $RLF$, the $QLF$ and the $PQLF$ loss functions. If the null is rejected, the model in the row is significantly better than model in the column. We only include in this table, the models which are not rejected as inferior by the $SPA$ test implemented with any of the three functions considered. When the $RLF$ is used, the test does not discriminate between the models when the significance level is 5%. However, when the significance level is 10%, the $TGARCH-N$ and the $GJR-N$ are superior to the $EGARCH-Skewed$, and the $EGARCH-Skewed$ is superior to the $APARCH-N$. With respect to the $QLF$ and $PQLF$, none of the models is rejected as inferior with the exception of the $EGARCH$ with $Skewed-t$ errors when compared with the $GJR$ with $GED$ errors by using the $QLF$[6].

Summarizing the results on estimating the $VaR$ of the $S\&P500$ returns considered in this section, it seems that the conclusion on the best specifications of the conditional variance and error distribution arising from the fitted models are those of the $TGARCH$, $EGARCH$ and $APARCH$ models which are similar between them. On the other hand, the estimates of the $VaR$ obtained when the conditional variances are specified by the $TGARCH$, $EGARCH$ and $APARCH$ models are almost identical. Furthermore, using bootstrap procedures to approximate the error distribution does not account for differences in the specification of the conditional variances. Therefore, we show that at least for the $S\&P500$ returns considered in this Chapter, it is more important to choose correctly the model for the conditional variance than to choose the error distribution. In any case, we may conclude that the asymmetry of the conditional distribution of returns seems to be more important than it having heavy tails. With respect to the backtesting procedures, we have seen that the test of Christoffersen (1998) is able to discriminate among the models but the $DQ$ and $LB$ tests only reject as inaccurate very few estimates of the $VaR$. When looking at the estimated $VaR$, the conclusions about which is the more appropriate model depend on the particular criteria chosen to compare them and the chosen loss function. Finally, the test proposed by Sarma et al. (2003) only discriminates very marginally.

### 2.4.3 Estimates of the $ES$

In this subsection, we estimate the $ES$ of the $S\&P500$ returns by a parametric, a bootstrap and a extreme value procedures. The parametric $ES^{\alpha}_{T+1}$ is calculated as the expected value of the returns beyond the quantile based on the assumed distribution. The Bootstrap $ES$ is calculated using the predictive distribution of returns, and the $EVT-ES$ is calculated using equations (2.29) and (2.30).

Figure 2.10 represents scatter-plots of the estimated $ES$ when the error distribution is $Normal$ and the expected value of the returns under the $VaR$ is estimated assuming

[6]The same results have been obtained when using the CPA test of Giacomini and White (2006).

this distribution. This figure is very similar to Figure 2.5 for the $VaR$ and shows that when the specification of the volatility is assumed to be $GARCH$, the estimates of the $ES$ are clearly different from the others. However, the estimated $ES$ obtained by the $TGARCH$, $EGARCH$ and $APARCH$ models are almost identical. The $GJR$ estimates are slightly different.

Figure 2.11 shows scatter-plots of the estimated $ES$ for the $EGARCH$ model using the seven different distributions considered. We can observe that the estimations of the $ES$ are very similar for all distributions. Therefore, comparing Figures 2.10 and 2.11, it seems that, as in the estimation of the $VaR$, an adequate specification of the variance is more important than the error distribution.

The two step procedure of Angelidis and Degiannakis (2007) is used for backtesting the $ES$. In the first step, the models that produce accurate $VaR$ forecast according to Christoffersen (1998) and reported in the previous subsection are selected. In the second step, the loss function is calculated using equation (2.31) and then the test of Hansen (2005) is used for evaluating the models. The corresponding p-values are reported in Table 2.9 where it can be observed that none of the models are rejected for being inferior than the others.

Finally, Table 2.10 reports the results of the test proposed by Sarma et al. (2003). In this case, the $GARCH-N$ is superior to the $GJR-N$ and the $GJR-N$ is superior to the $APARCH-N$ at the 10% confidence level. However, the $APARCH-N$ is not rejected when compared with the $GARCH-N$ and, consequently the results are somehow ambiguous.

## 2.5 Conclusions

In this Chapter, we update previous reviews of the literature on estimating and backtesting the $VaR$ and $ES$ measures of financial risk. According to the requirements of the Basel Committee, the risk measures are calculated at the 1% confidence level instead of the 5% and 10% levels usually found in the literature. This distinction may have important consequences as the difficulty involved in estimating quantiles increases as the confidence decreases. The different estimators and tests have been illustrated by estimating the $VaR$ and $ES$ of a series of daily returns of the $S\&P500$ index. In this case, we first conclude that it seems more important to choose an adequate model for the conditional variance than to choose correctly the distribution of the errors. It seems that modeling the leverage effect is important when predicting the $VaR$ and the $ES$. However, different ways of computing the quantile of the distribution of standardized returns only have marginal effects on whether a model is rejected or not. Furthermore, comparing the $VaR$ and $ES$ estimates obtained by the different models, we also observe that the differences among them are larger when fixing the distribution and comparing the models than when the model is fixed and the distribution changes. The importance of leverage effect when predicting the $VaR$ was already pointed out by Engle (2003) in his Nobel lecture. The particular model chosen as the best depends on the criteria chosen for the selection but it seems that models with leverage effect and Skewed-t distributions are more preferred. In any case, the $HS$ and Adaptive $CAViaR$ estimates are too smooth and rejected as adequate. On the other hand, the Asymmetric $CAViaR$ estimates are not rejected and less conservative when compared with the estimates obtained by alternative models. Our empirical results also suggest that using basic bootstrap procedures to estimate the quantile of the error distribution do not account for the potential misspecification of the conditional variance.

It is of interest to analyze whether the importance of choosing an adequate specification for the conditional variance as opposite to choosing a correct distribution of the errors can be generalized to other time series of returns.

We also find that Christoffersen (1998) test is the only one able to discriminate among alternative $VaR$ estimates. There is not a clear candidate test to choose among accurate estimates of the $VaR$. Finally, there are very few proposals for backtesting $ES$ estimates.

Our results are somehow different from those of Jalal and Rockinger (2008) who conclude that the $GPD$ estimator of McNeil and Frey (2000) has a good performance in front of misspecification when the data is generated by $GARCH$ models with leverage effect, regime switching model or Stochastic Volatility models with jumps but the symmetric $GARCH$ model is fitted. According to our results it is very important to choose an adequate specification of the conditional variance in order to have adequate estimates of the $VaR$ and $ES$.

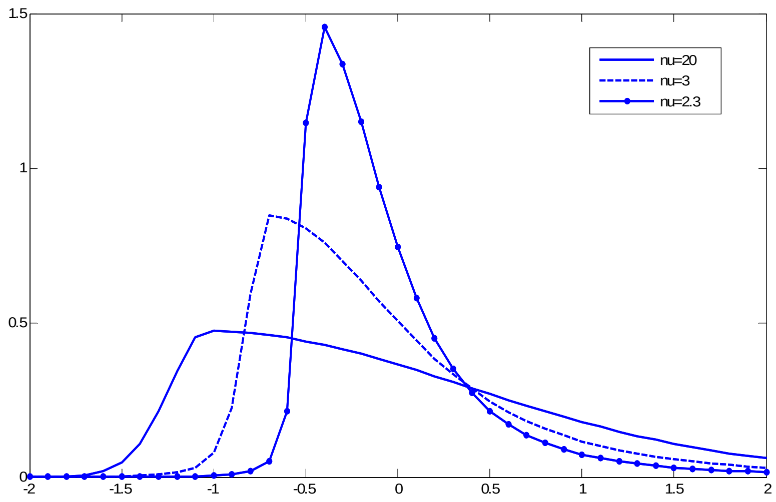
Another interesting topic for further research is the extension of the analysis carried out in this Chapter to incorporate Stochastic Volatility ($SV$) models in the comparison.

Some applications of $SV$ models to the estimation of the $VaR$ can be found in Billio and Pelizzon (2000), Billio and Sartore (2003), Eberlein et al. (2003) and Sadorsky (2005); see the considerations of McAleer (2009) on the different alternatives that should be taken into account when estimating the $VaR$.

Also, it would also be interesting to incorporate in the comparison the estimates of the $VaR$ based on ultra-high-frequency data volatility measures; see Brownless and Gallo (2010) for estimates of the $VaR$ based on measures of volatility related to realized volatility.

Finally, it is also interesting to compare the alternative models by evaluating the entire forecast density as proposed by Berkowitz (2001).

Figure 2.1. Skewed Student density function for different degrees of freedom and asymmetry parameter: $(a)$ $\xi = 0.75$ and $(b)$ $\xi = -0.75$.



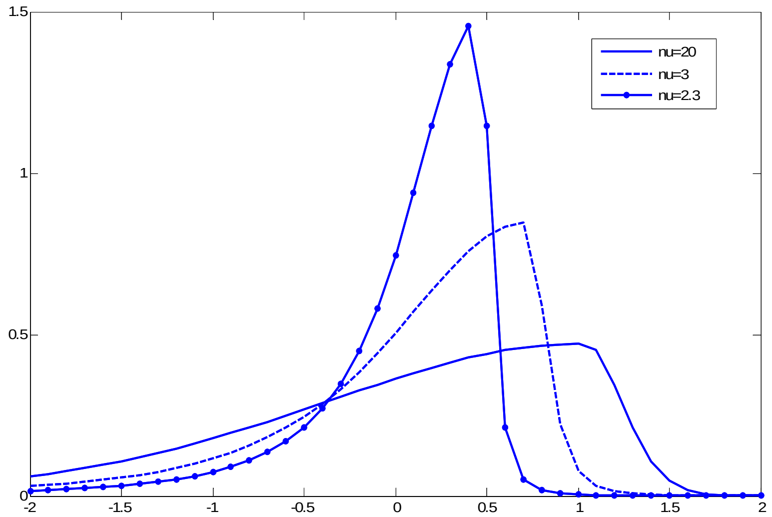



Figure 2.2. Hill and GPD estimators of $q_{0.01}$ for SP&500 returns as a function of the number of observations in the tail.

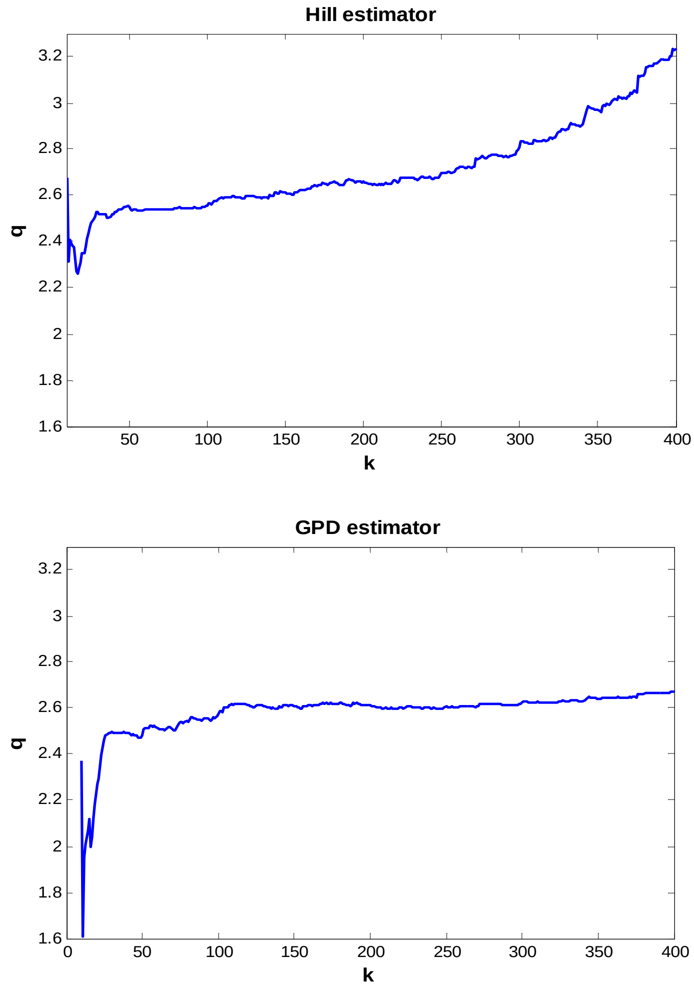


Table 2.1. Summary of the models proposed to estimate the $VaR$

| References | Conditional Mean | Conditional Variance | Distribution of $\epsilon_t$ |
|---|---|---|---|
| Morgan (1995) | | $IGARCH\,(1,1)$ | Normal |
| Barone-Adesi et al. (1999) | $ARMA\,(1,1)$ | $AGARCH(1,1)$ | Bootstrap |
| Billio and Pelizzon (2000) | | $SV$ | Switching Regime |
| Danielsson and de Vries (2000) | | $GARCH(1,1)$ | EVT-Hill |
| McNeil and Frey (2000) | $AR\,(1)$ | $GARCH\,(1,1)$ | EVT-GDP |
| Mittnik and Paolella (2000) | $AR\,(1)$ | $GARCH(1,1)$, $APARCH\,(1,1)$ | Normal, Student |
| Barone-Adesi and Giannopoulos (2001) | $MA\,(1)$ | $AGARCH(1,1)$ | Bootstrap |
| Nystrom and Skoglund (2002) | $ARMA(p,q)$ | $GJRGARCH(1,1)$ | EVT-GDP, EVT-Hill |
| Billio and Sartore (2003) | | $SV$ | Bootstrap |
| Eberlein et al. (2003) | | $SV$ | Hyperbolic |
| Giot and Laurent (2003) | $AR\,(p)$ | $IGARCH\,(1,1)$ , $APARCH\,(1,1)$ | Normal, Student, Skewed-t |
| Angelidis et al. (2005) | $AR\,(1)$ | $GARCH(1,1)$, $GJRGARCH(1,1)$, $EGARCH(1,1)$ | Normal, Student, GED |
| Christoffersen and GonÇalves (2005) | | $GARCH\,(1,1)$ | Normal, Student, EVT-Hill, Bootstrap |
| Giannopoulos and Tunaru (2005) | $MA\,(1)$ | $AGARCH(1,1)$ | Bootstrap |
| Sadorsky (2005) | | $SV$ | Normal |
| Hartz and Paolella (2006) | $ARMA(p,q)$ | $GARCH\,(1,1)$ | Bootstrap |
| Bali and Theodossiou (2007) | $AR\,(1)$ | All $GARCH$ models included in (12) | Normal, SGT, GED |
| Engle and Manganelli (2004) | Conditional Quantile | | |
| Chen and Chen (2005) | Dynamic Adaptive Quantile | | |
| Gourieroux and Jasiak (2006) | Quantile Regression | | |

Table 2.2. Summary of the models proposed to estimate the $ES$

| References | Conditional Mean | Conditional Variance | Distribution of $\epsilon_t$ |
|---|---|---|---|
| McNeil and Frey (2000) | $AR\,(1)$ | $GARCH\,(1,1)$ | EVT-GDP |
| Giot and Laurent (2003) | $AR\,(p)$ | $IGARCH\,(1,1)$ , $APARCH\,(1,1)$ | Historical Simulation |
| Christoffersen and GonÇalves (2005) | | $GARCH\,(1,1)$ | Normal, Student, EVT-Hill |
| Giannopoulos and Tunaru (2005) | $MA\,(1)$ | $AGARCH(1,1)$ | Bootstrap |
| Bali and Theodossiou (2007) | $AR\,(1)$ | All $GARCH$ models included in (12) | Historical Simulation |

Figure 2.3. Daily S&P500 returns observed from 29th August 1995 up to 20th October 2005, correlograms of returns and absolute returns and cross-correlogram of returns and squared returns together with their corresponding 95% confidence bands.

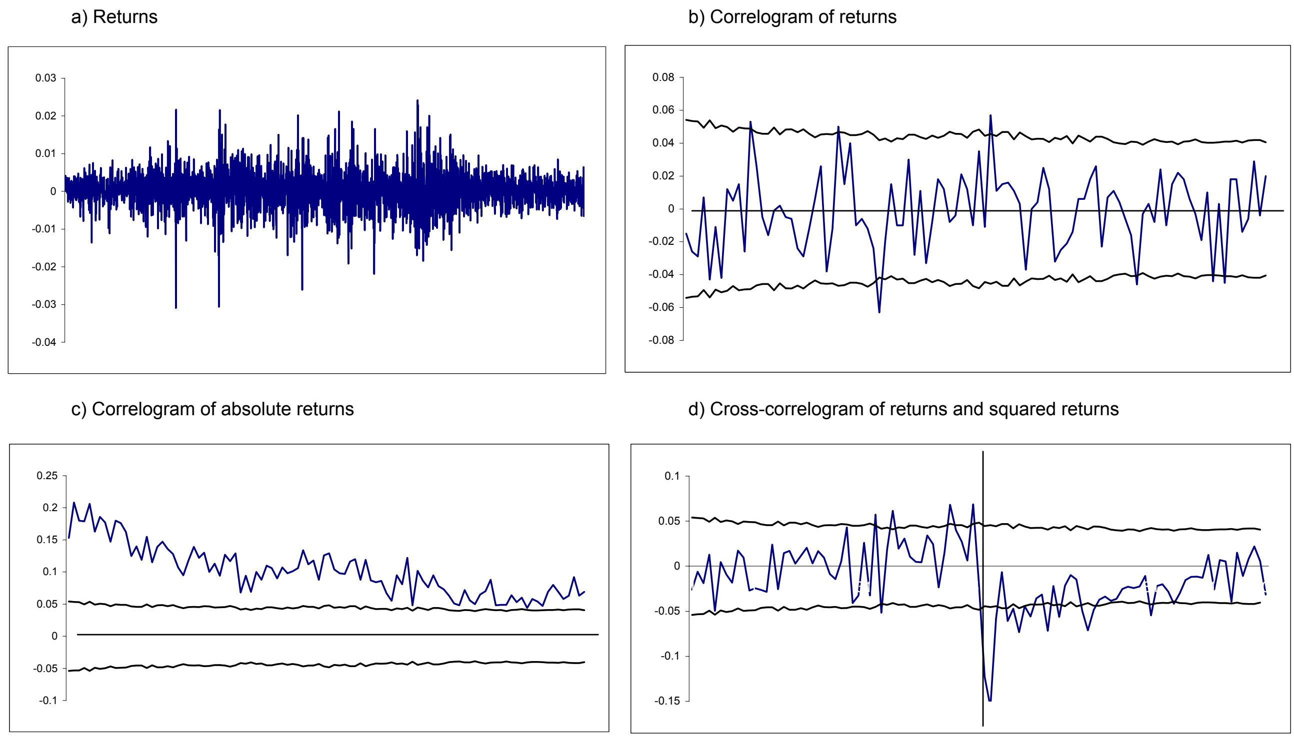


Table 2.3. Descriptive statistics of daily S&P500 returns observed from 29th August 1995 up to 20th October 2005. The quantities reported for $Q(20)$ and $Q^*_{[M/3]-1}(20)$ are the ratios between the value of the statistic and its corresponding 5% critical value.

| | |
|---|---|
| Sample size | 2555 |
| Mean | 0.0001 |
| Median | 0.0002 |
| Maximum | 0.0242 |
| Minimum | $-0.0309$ |
| Std. Dev. | 0.0050 |
| Skewness | $-0.0971$* |
| Kurtosis | 6.0215* |
| $Q(20)$ | 0.90 |
| $Q^*_{[M/3]-1}(20)$ | 389.06* |
| $Corr\left(y_t, y^2_{t+1}\right)$ | $-0.1225$* |

* Significant values at 5% level.

Figure 2.4. Kernel densities of standardized residuals obtained using the $GARCH$ (first row) and $APARCH$ (second row) specifications of the conditional variances with different assumptions of the error distribution (continuous line) together with the estimated density (discontinuous line).

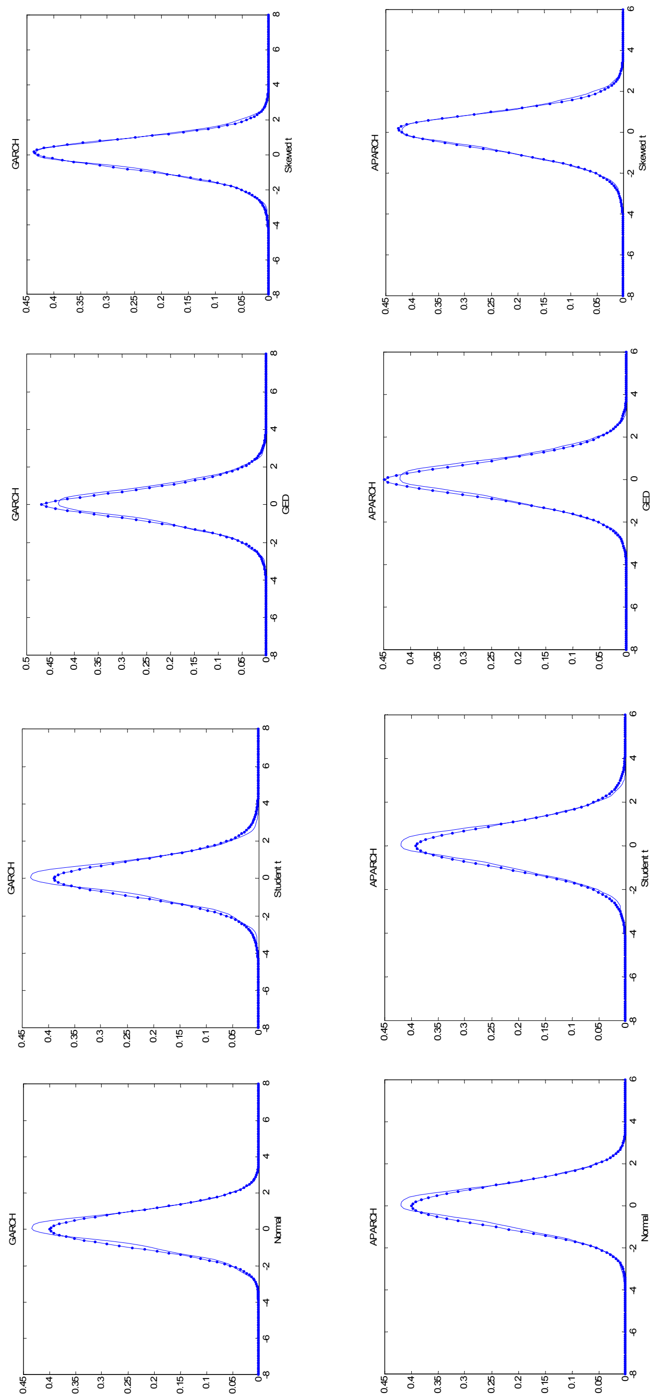


Table 2.4. Maximum Likelihood estimates of parameters of the alternative GARCH-type models with different conditional distributions fitted to daily S&P500 returns.

| Model | | Normal | Student-$\nu$ | GED | Skewed-t |
|---|---|---|---|---|---|
| GARCH | $\alpha_0$ | $2.49E-07$ ($8.30E-08$) | $2.10E-07$ ($5.78E-08$) | $1.96E-07$ ($5.78E-08$) | $2.15E-07$ ($1.17E-07$) |
| | $\alpha_1$ | 0.080 (0.014) | 0.066 (0.011) | 0.071 (0.013) | 0.072 (0.021) |
| | $\beta_1$ | 0.912 (0.013) | 0.925 (0.011) | 0.922 (0.012) | 0.921 (0.0216) |
| | $\nu$ | | 9.881 (1.993) | 1.535 (0.080) | 10.236 (2.080) |
| | $\xi$ | | | | $-0.117$ (0.025) |
| TGARCH | $\alpha_0$ | $1.10E-04$ ($3.22E-05$) | $9.99E-05$ ($2.43E-05$) | $1.05E-04$ ($2.93E-05$) | $9.77E-05$ ($4.63E-05$) |
| | $\alpha_1$ | 0.011 (0.011) | 0.007 (0.028) | 0.009 (0.005) | 0.010 (0.020) |
| | $\beta_1$ | 0.924 (0.015) | 0.931 (0.018) | 0.927 (0.011) | 0.930 (0.019) |
| | $\gamma$ | $-0.115$ (0.021) | $-0.110$ (0.021) | $-0.113$ (0.020) | $-0.108$ (0.032) |
| | $\nu$ | | 12.705 (3.289) | 1.635 (0.085) | 13.340 (3.598) |
| | $\xi$ | | | | $-0.110$ (0.026) |
| GJR | $\alpha_0$ | $4.07E-07$ ($1.02E-07$) | $3.31E-07$ ($6.23E-08$) | $3.57E-07$ ($7.44E-08$) | $3.32E-07$ ($1.13E-07$) |
| | $\alpha_1$ | 0.002 (0.019) | 0.000 (0.016) | 0.000 (0.005) | 0.000 (0.008) |
| | $\beta_1$ | 0.913 (0.023) | 0.924 (0.014) | 0.921 (0.011) | 0.924 (0.016) |
| | $\gamma$ | 0.140 (0.028) | 0.128 (0.022) | 0.133 (0.023) | 0.128 (0.034) |
| | $\nu$ | | 12.160 (3.041) | 1.620 (0.088) | 13.237 (3.730) |
| | $\xi$ | | | | $-0.120$ (0.026) |
| EGARCH | $\alpha_0$ | $-0.351$ (0.084) | $-0.323$ (0.068) | $-0.337$ (0.065) | $-0.320$ (0.115) |
| | $\alpha_1$ | 0.128 (0.021) | 0.119 (0.0168) | 0.123 (0.0179) | 0.122 (0.029) |
| | $\beta_1$ | 0.976 (0.006) | 0.978 (0.005) | 0.977 (0.0052) | 0.978 (0.009) |
| | $\gamma$ | $-0.102$ (0.018) | $-0.099$ (0.015) | $-0.101$ (0.0151) | $-0.097$ (0.026) |
| | $\nu$ | | 12.909 (3.425) | 1.638 (0.086) | 13.608 (3.660) |
| | $\xi$ | | | | $-0.112$ (0.026) |
| APARCH | $\alpha_0$ | $6.55E-05$ ($7.788E-05$) | $4.12E-05$ ($6.68E-05$) | $1.10E-04$ ($3.50E-05$) | $1.95E-05$ ($4.02E-05$) |
| | $\alpha_1$ | 0.067 (0.013) | 0.059 (0.011) | 0.066 (0.010) | 0.058 (0.019) |
| | $\beta_1$ | 0.923 (0.014) | 0.929 (0.011) | 0.925 (0.014) | 0.928 (0.020) |
| | $\gamma$ | 0.834 (0.118) | 0.893 (0.151) | 0.872 (0.114) | 0.847 (0.287) |
| | $\nu$ | | 12.705 (3.287) | 1.635 (0.084) | 13.410 (3.658) |
| | $\xi$ | | | | $-0.114$ (0.027) |
| | $\delta$ | 1.093 (0.202) | 1.159 (0.282) | 0.999 (0.084) | 1.287 (0.363) |

The quantities in parenthesis are asymptotic standard errors.

Table 2.5. Diagnostics on the standardized residuals of GARCH models.

| | | *Skewness* | *Kurtosis* | $Q(20)$ | $Q^*_{[20/3]-1}(20)$ | $Corr\left(y_t, y^2_{t+1}\right)$ |
|---|---|---|---|---|---|---|
| GARCH | Normal | −0.4253 (0) | 4.5516* (3) | 0.6959 | 0.4613 | −0.0953* |
| | Student-$\nu$ | −0.4284* (0) | 4.6206* (4.02) | 0.6424 | 0.8324 | −0.098* |
| | GED | −0.4293* (0) | 4.5932* (3.68) | 0.6440 | 0.6494 | −0.097* |
| | Skewed-t | −0.4286* (−0.2803) | 4.5910* (4.03) | 0.6456 | 0.6301 | −0.0969* |
| TGARCH | Normal | −0.3460* (0) | 4.0828* (3) | 0.7229 | 0.5275 | −0.0583* |
| | Student-$\nu$ | −0.3500* (0) | 4.1137* (3.68) | 0.7193 | 0.6897 | −0.0612* |
| | GED | −0.3478* (0) | 4.0969* (3.48) | 0.7213 | 0.6098 | −0.0594* |
| | Skewed-t | −0.3519* (−0.2365) | 4.1118* (3.69) | 0.7185 | 0.5267 | −0.0622* |
| GJR | Normal | −0.3883* (0) | 4.2497* (3) | 0.7480 | 0.3278 | −0.0592* |
| | Student-$\nu$ | −0.3993* (0) | 4.3196* (3.73) | 0.7421 | 0.3817 | −0.0637* |
| | GED | −0.3957* (0) | 4.2970* (3.51) | 0.7446 | 0.3597 | −0.062* |
| | Skewed-t | −0.3994* (−0.2589) | 4.3204* (3.70) | 0.7420 | 0.3817 | −0.0638* |
| EGARCH | Normal | −0.3476* (0) | 4.0723* (3) | 0.7284 | 0.4406 | −0.0595* |
| | Student-$\nu$ | −0.3509* (0) | 4.0960* (3.67) | 0.7245 | 0.5164 | −0.0619* |
| | GED | −0.3491* (0) | 4.0838* (3.48) | 0.7266 | 0.4638 | −0.0604* |
| | Skewed-t | −0.3526* (−0.2389) | 4.0945* (3.67) | 0.7240 | 0.3984 | −0.0627* |
| APARCH | Normal | −0.3504* (0) | 4.0951* (3) | 0.7298 | 0.4923 | −0.0581* |
| | Student-$\nu$ | −0.3583* (0) | 4.1409* (3.68) | 0.7302 | 0.6630 | −0.0609* |
| | GED | −0.3451* (0) | 4.0881* (3.48) | 0.7229 | 0.6695 | −0.0579* |
| | Skewed-t | −0.3666* (−0.2445) | 4.1646* (3.69) | 0.7352 | 0.4652 | −0.0619* |

* Significant values at 5% level

The figures in parenthesis in the column of skewness and kurtosis, represents the corresponding population moments implied by the estimated distribution.

Table 2.6. p-values of Backtesting $VaR1\%$ tests for S&P500 based on n=1000 days.

| | | Christoffersen Likelihood | DQ Test | LB(5) | LB(20) |
|---|---|---|---|---|---|
| | | p-value | p-value | p-value | p-value |
| Historical Simulation | | 0.031 | 0.004 | 7.33E-05 | 5.48E-29 |
| CAViaR Adaptive | | 0.008 | 0.349 | 0.999 | 0.999 |
| CAViaR Asymmetric | | 0.375 | 0.888 | 0.999 | 6.23E-05 |
| GARCH | Normal | 0.57 | 0.984 | 0.999 | 0.999 |
| | Student- v | 0.008 | 0.372 | 0.999 | 0.999 |
| | GED | 0.032 | 0.548 | 0.999 | 0.999 |
| | Skewed-t | 0.001 | 0.228 | 0.999 | 0.999 |
| | Bootstrap | 0.374 | 0.932 | 0.999 | 0.999 |
| | EVT-Hill | 0.008 | 0.369 | 0.999 | 0.999 |
| | EVT-GPD | 0.008 | 0.369 | 0.999 | 0.999 |
| TGARCH | Normal | 0.208 | 0.862 | 0.999 | 0.999 |
| | Student- v | 0.093 | 0.718 | 0.999 | 0.999 |
| | GED | 0.093 | 0.721 | 0.999 | 0.999 |
| | Skewed-t | 0.093 | 0.716 | 0.999 | 0.999 |
| | Bootstrap | 0.094 | 0.727 | 0.999 | 0.999 |
| | EVT-Hill | 0.094 | 0.721 | 0.999 | 0.999 |
| | EVT-GPD | 0.094 | 0.721 | 0.999 | 0.999 |
| GJR | Normal | 0.207 | 0.862 | 0.999 | 0.999 |
| | Student- v | 0.207 | 0.855 | 0.999 | 0.999 |
| | GED | 0.207 | 0.859 | 0.999 | 0.999 |
| | Skewed-t | 0.094 | 0.709 | 0.999 | 0.999 |
| | Bootstrap | 0.208 | 0.864 | 0.999 | 0.999 |
| | EVT-Hill | 0.094 | 0.718 | 0.999 | 0.999 |
| | EVT-GPD | 0.094 | 0.718 | 0.999 | 0.999 |
| EGARCH | Normal | 0.094 | 0.722 | 0.999 | 0.999 |
| | Student- v | 0.094 | 0.718 | 0.999 | 0.999 |
| | GED | 0.094 | 0.720 | 0.999 | 0.999 |
| | Skewed-t | 0.375 | 0.717 | 0.999 | 0.999 |
| | Bootstrap | 0.094 | 0.724 | 0.999 | 0.999 |
| | EVT-Hill | 0.094 | 0.721 | 0.999 | 0.999 |
| | EVT-GPD | 0.094 | 0.721 | 0.999 | 0.999 |
| APARCH | Normal | 0.208 | 0.861 | 0.999 | 0.999 |
| | Student- v | 0.033 | 0.543 | 0.999 | 0.999 |
| | GED | 0.094 | 0.720 | 0.999 | 0.999 |
| | Skewed-t | 0.208 | 0.712 | 0.999 | 0.999 |
| | Bootstrap | 0.094 | 0.727 | 0.999 | 0.999 |
| | EVT-Hill | 0.094 | 0.720 | 0.999 | 0.999 |
| | EVT-GPD | 0.094 | 0.720 | 0.999 | 0.999 |

The shaded areas represent models that are not rejected when $\alpha$=0.10 under the different backtesting methods.

Figure 2.5. Scatter-plots of the $VaR1\%$ estimated by the alternative conditionally heteroscedastic models with $Normal$ errors.

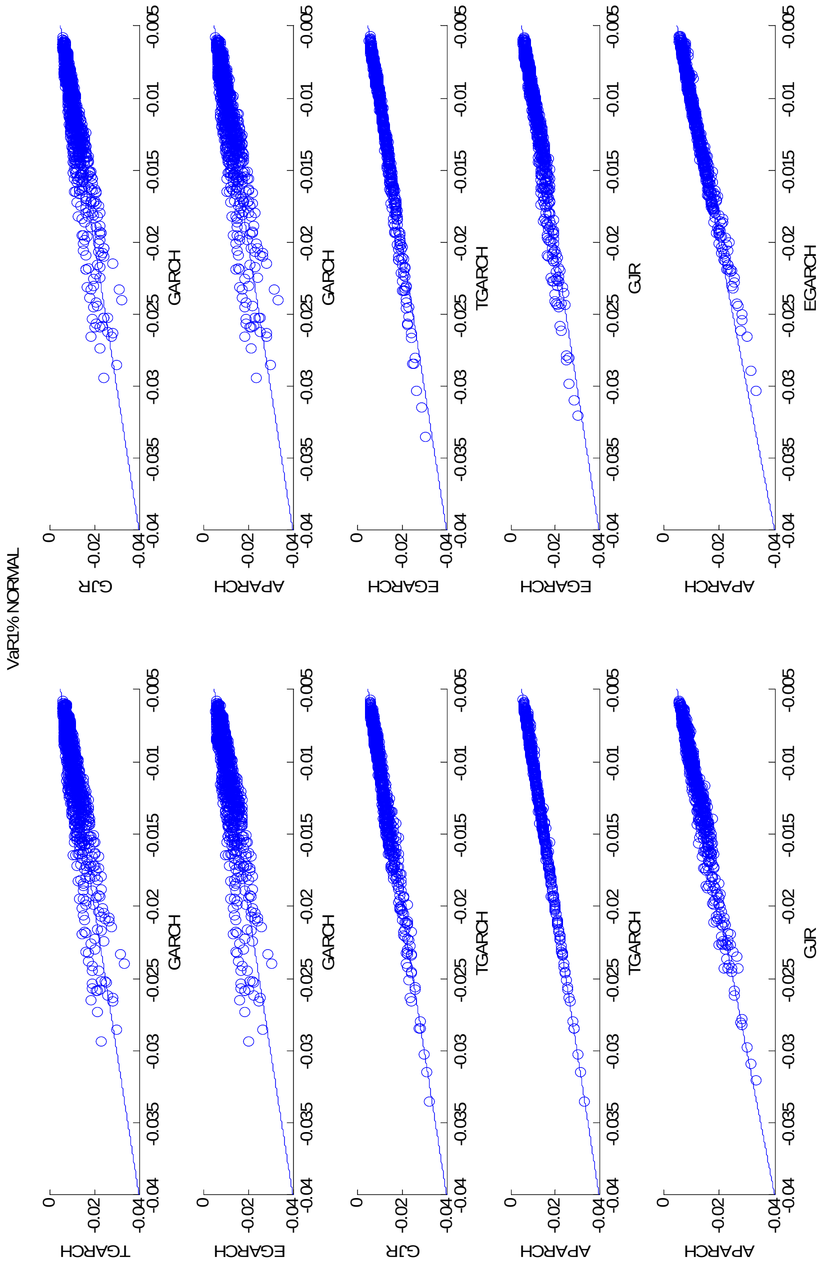


Figure 2.6. Scatter-plots of the $VaR$ estimated by the alternative conditionally heteroscedastic models with *Bootstrap* errors.

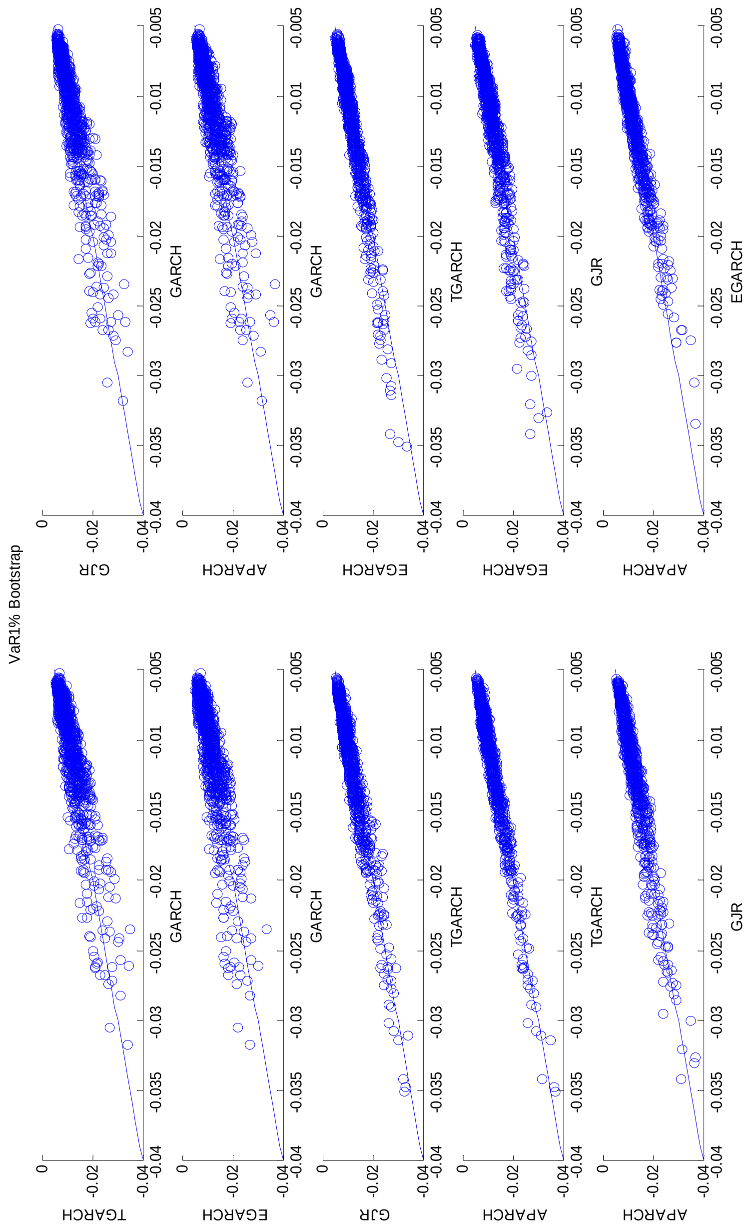


Figure 2.7. Scatter-plots of the $VaR1\%$ estimated by the alternative error distributions with the $EGARCH$ model.

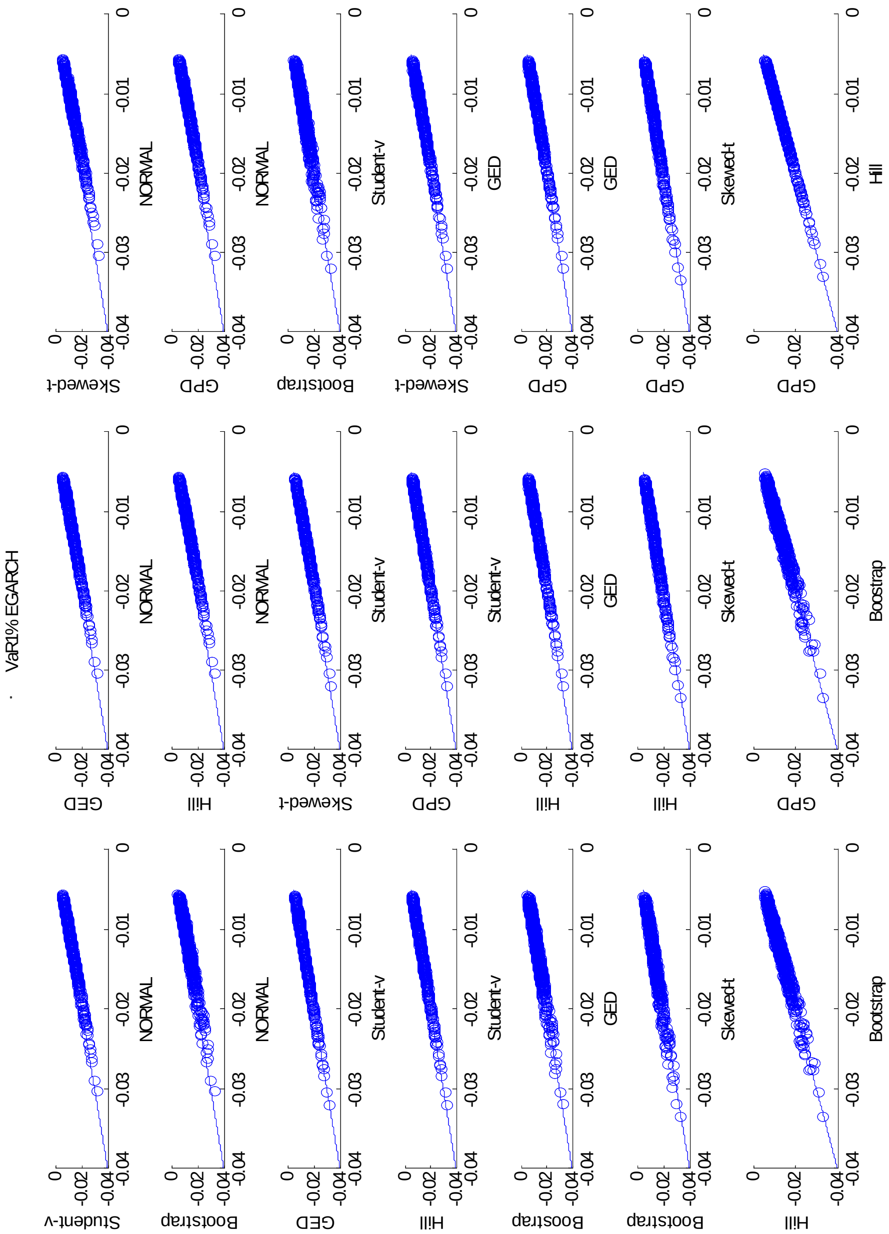


Figure 2.8. Scatter-plots of the $VaR1\%$ estimated with the $CAViaR\ Asymmetric$ and the $EGARCH$ model with different error distributions.

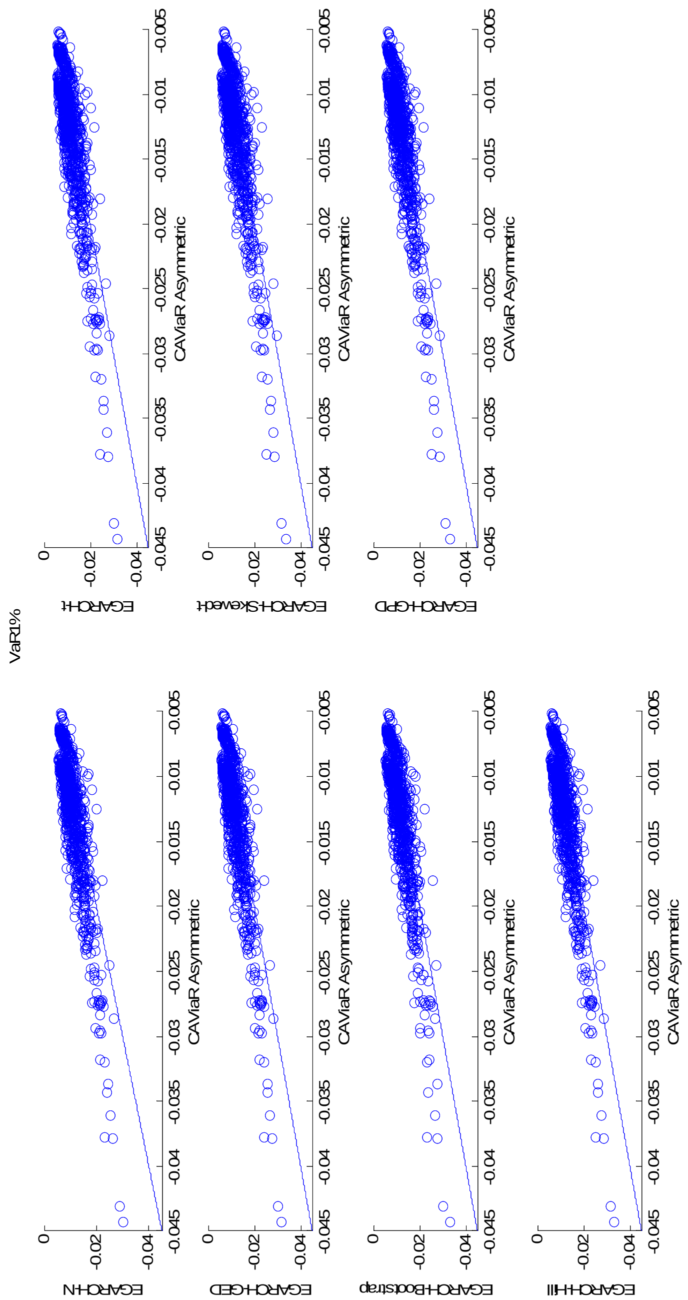


Figure 2.9. Estimates of the $VaR1\%$ obtained using the $CAViaR$ and $HS$ procedures and the $EGARCH$ model with different error distributions.

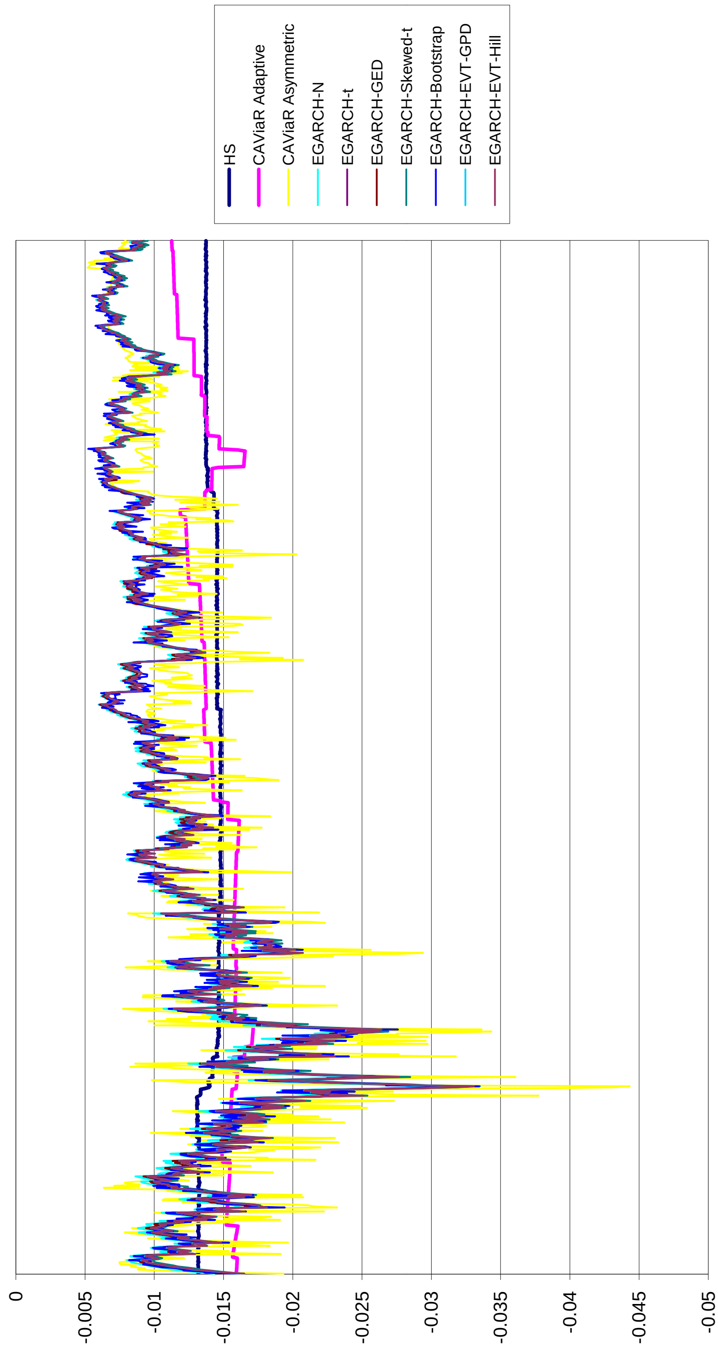


Table 2.7. Test proposed by Lopez (1999) and SPA test to compare accurate $VaR1\%$ estimates computed with alternative loss functions.

| | Loss Function | | | SPA test | | |
|---|---|---|---|---|---|---|
| Model | RLF | QLF | PQL | RLF | QLF | PQLF |
| | | | | p-value | p-value | p-value |
| $CAViaR\ Asymmetric$ | 0.6567 | 0.0261 | 0.1423 | 0.5840 | 0.1600 | 0.0010 |
| $GARCH-N$ | 0.2494 | 0.0214 | 0.1170 | 0.9820 | 0.9920 | 0.9990 |
| $GARCH-Bootstrap$ | 0.3041 | 0.0235 | 0.1181 | 0.6790 | 0.0001 | 0.9510 |
| $TGARCH-N$ | 0.3671 | 0.0214 | 0.1209 | 0.3530 | 0.9150 | 0.7450 |
| $GJR-N$ | 0.3838 | 0.0221 | 0.1201 | 0.2190 | 0.4790 | 0.8750 |
| $GJR-t$ | 0.1871 | 0.0227 | 0.1204 | 0.9890 | 0.0540 | 0.8240 |
| $GJR-GED$ | 0.2242 | 0.0223 | 0.1214 | 0.4650 | 0.2540 | 0.2280 |
| $GJR-Bootstrap$ | 0.3001 | 0.0244 | 0.1254 | 0.3080 | 0.0020 | 0.0001 |
| $EGARCH-Skewed-t$ | 0.1584 | 0.0211 | 0.1242 | 0.9980 | 0.9930 | 0.1330 |
| $APARCH-N$ | 0.3600 | 0.0215 | 0.1206 | 0.3410 | 0.9100 | 0.8910 |
| $APARCH-Skewed-t$ | 0.3267 | 0.0236 | 0.1239 | 0.9999 | 0.0790 | 0.2170 |

Table 2.8. Test proposed by Sarma et al. (2003) to compare accurate estimates of $VaR1\%$ computed with alternative loss functions.

| $VaR$ Regulatory Loss Function | | | | | |
|---|---|---|---|---|---|
| Model | $TGARCH-N$ | $GJR-N$ | $GJR-GED$ | $EGARCH-Skewed-t$ | $APARCH-N$ |
| $GARCH-N$ | −0.5251 (0.5996) | −0.7895 (0.4300) | 0.1666 (0.8677) | 0.4385 (0.6611) | −0.4924 (0.6225) |
| $TGARCH-N$ | | −0.2512 (0.8016) | 1.3925 (0.1641) | 1.8655 (0.0624) | 1.1640 (0.2447) |
| $GJR-N$ | | | 1.8995 (0.0578) | 1.7872 (0.0742) | 0.3514 (0.7253) |
| $GJR-GED$ | | | | 1.0487 (0.2945) | −1.3207 (0.1869) |
| $EGARCH-Skewed-t$ | | | | | −1.8191 (0.0692) |

| $VaR$ Quantile Loss Function | | | | | |
|---|---|---|---|---|---|
| Model | $TGARCH-N$ | $GJR-N$ | $GJR-GED$ | $EGARCH-Skewed-t$ | $APARCH-N$ |
| $GARCH-N$ | −0.0664 (0.9470) | −0.5592 (0.5761) | −0.6466 (0.5180) | 0.1850 (0.8532) | −0.0708 (0.9435) |
| $TGARCH-N$ | | −1.3694 (0.1712) | −1.1860 (0.2359) | 0.7253 (0.4684) | −0.0497 (0.9604) |
| $GJR-N$ | | | −0.3517 (0.7251) | 1.3219 (0.1865) | 1.3282 (0.1844) |
| $GJR-GED$ | | | | 1.8038 (0.0716) | 1.1516 (0.2497) |
| $EGARCH-Skewed-t$ | | | | | −0.6980 (0.4853) |

| $VaR$ Predictive Quantile Loss Function | | | | | |
|---|---|---|---|---|---|
| Model | $TGARCH-N$ | $GJR-N$ | $GJR-GED$ | $EGARCH-Skewed-t$ | $APARCH-N$ |
| $GARCH-N$ | −0.7729 (0.4397) | −0.7018 (0.4829) | −1.0829 (0.2791) | −1.3709 (0.1707) | −0.7285 (0.4665) |
| $TGARCH-N$ | | 0.5994 (0.5490) | −0.3583 (0.7202) | −1.2798 (0.2009) | 1.0034 (0.3159) |
| $GJR-N$ | | | −0.7977 (0.4252) | −1.2178 (0.2236) | −0.3760 (0.7070) |
| $GJR-GED$ | | | | −1.3424 (0.1797) | 0.5325 (0.5944) |
| $EGARCH-Skewed-t$ | | | | | 1.4093 (0.1590) |

*p-values in parentheses

Figure 2.10. Scatter-plots of the *ES1%* estimated by the alternative conditionally heteroscedastic models with *Normal* errors.

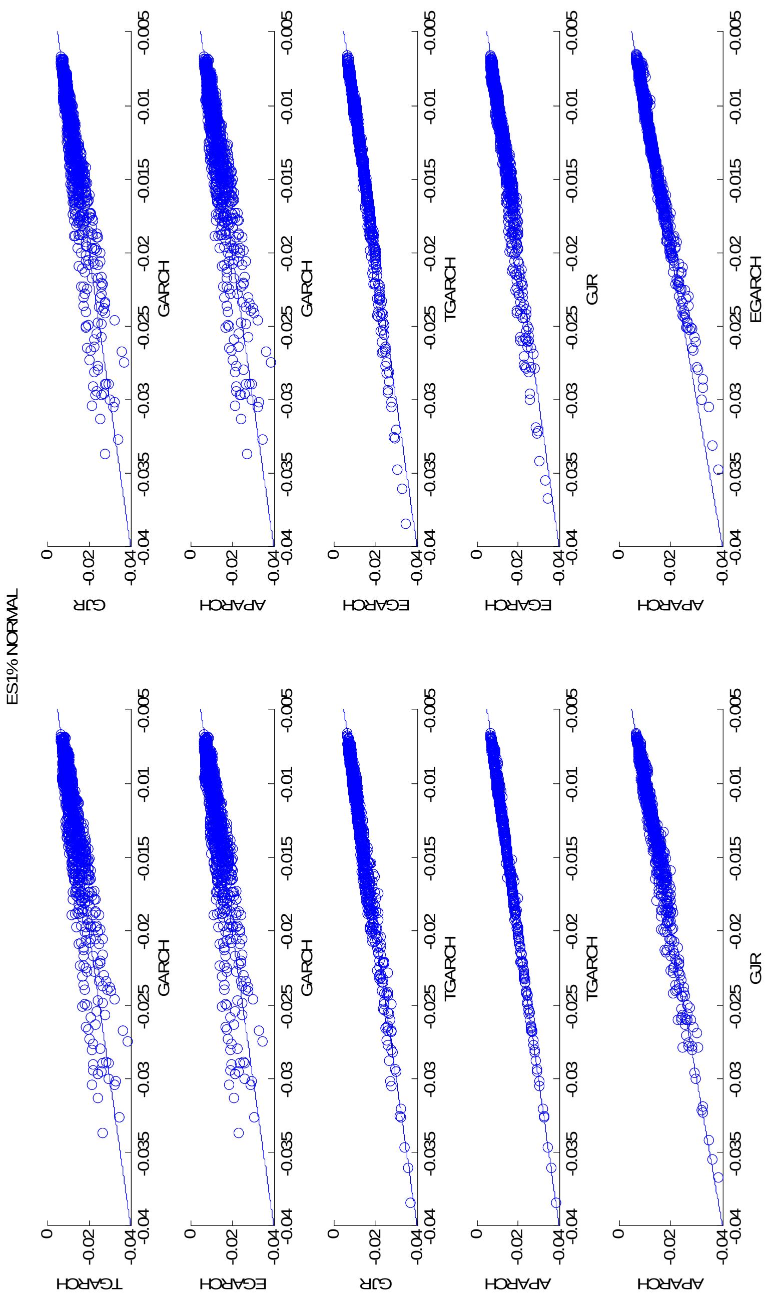


Figure 2.11. Scatter-plots of the $ES1\%$ estimated by the alternative error distributions with the $EGARCH$ model.

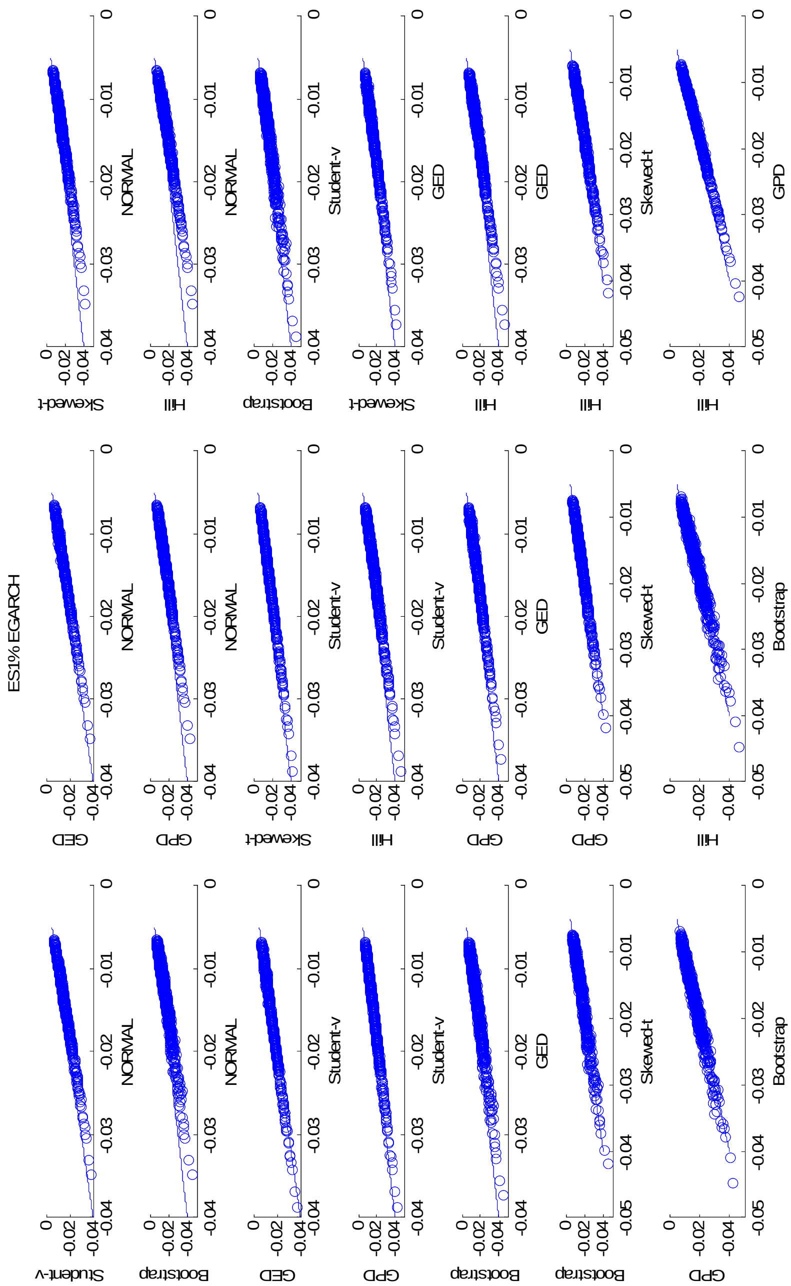


Table 2.9. SPA test to compare $ES1\%$ estimates obtained from accurate $VaR$ estimates.

| Model | SPA test RLF p-value |
|---|---|
| $CAViaR\ Asymmetric$ | 0.7430 |
| $GARCH-N$ | 0.2920 |
| $GARCH-Bootstrap$ | 0.7680 |
| $TGARCH-N$ | 0.2470 |
| $GJR-N$ | 0.9999 |
| $GJR-t$ | 0.4520 |
| $GJR-GED$ | 0.3940 |
| $GJR-Bootstrap$ | 0.4460 |
| $EGARCH-Skewed-t$ | 0.6530 |
| $APARCH-N$ | 0.2660 |
| $APARCH-Skewed-t$ | 0.6840 |

Table 2.10. Test proposed by Sarma et al. (2003) to compare accurate estimates of VaR using Angelidis and Degiannakis (2007) loss function.

| | $ES$ Regulatory Loss Function | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Model | $GARCH-N$ | $GARCH-Bootstrap$ | $TGARCH-N$ | $GJR-N$ | $GJR-t$ | $GJR-GED$ | $GJR-Bootstrap$ | $EGARCH-Skewed-t$ | $APARCH-N$ | $APARCH-Skewed-t$ |
| $CAViaR\ Asymmetric$ | −0.3282 (0.7428) | 0.8440 (0.3988) | 0.8207 (0.4120) | 1.1650 (0.2443) | 0.7302 (0.4654) | 0.7459 (0.4558) | 0.8535 (0.3936) | −0.0866 (0.9309) | 0.7586 (0.4482) | −0.8466 (0.3974) |
| $GARCH-N$ | | 1.6178 (0.1060) | 1.3219 (0.1865) | 1.9131 (0.0560) | 1.1097 (0.2674) | 1.1303 (0.2586) | 1.3367 (0.1816) | 0.2900 (0.7718) | 1.2594 (0.2082) | −0.1771 (0.8595) |
| $GARCH-Bootstrap$ | | | −0.4704 (0.6381) | 1.0113 (0.3121) | −0.7091 (0.4784) | −0.6794 (0.4970) | −0.4431 (0.6578) | −0.8750 (0.3818) | −0.5568 (0.5778) | −0.9964 (0.3193) |
| $TGARCH-N$ | | | | 1.6406 (0.1012) | −1.1233 (0.2615) | −1.0232 (0.3064) | −0.0343 (0.9726) | −0.8764 (0.3810) | −1.6336 (0.1027) | −1.0388 (0.2991) |
| $GJR-N$ | | | | | −1.5766 (0.1152) | −1.5407 (0.1237) | −1.4615 (0.1442) | −1.1991 (0.2308) | −1.7076 (0.0880) | −1.2202 (0.2227) |
| $GJR-t$ | | | | | | 1.0030 (0.3161) | 1.3142 (0.1891) | −0.8107 (0.4177) | 0.8080 (0.4193) | −1.0260 (0.3051) |
| $GJR-GED$ | | | | | | | 1.2120 (0.2258) | −0.8336 (0.4047) | 0.7094 (0.4782) | −1.0403 (0.2984) |
| $GJR-Bootstrap$ | | | | | | | | −0.8980 (0.3694) | −0.5311 (0.5954) | −1.0657 (0.2868) |
| $EGARCH-Skewed-t$ | | | | | | | | | 0.8204 (0.4121) | −1.1515 (0.2498) |
| $APARCH-N$ | | | | | | | | | | −1.0014 (0.3168) |

*p-values in parentheses

# Chapter 3

# Robustness against VaR and ES level and horizon

## 3.1 Introduction

In order to analyze the robustness of the results obtained in the previous Chapter, we need to calculate the $VaR$ and the $ES$ at different points of the tail of the distribution of returns, in particular, for the 5% and the 10% worst cases. We backtest the procedures and choose the best models in each case and compare with the ones obtained for the 1%. Furthermore, we graphically analyze the importance of the specification of the conditional variance and the distribution of returns.

Another issue related to regulation of the financial institutions is the forecast horizon. As we mention before, for determining minimal capital requirements, regulators ask for $VaR$ estimates for 10 days returns. However, the industry, with the purpose of internal risk control, has all the infrastructure for the estimation of daily $VaR$. Therefore, for reporting the ten steps-ahead $VaR$, the financial institutions use, for simplicity, the square root of time. There are very few papers in the literature devoted to propose alternatives to the square root of time. Thus, in section 3.2 we compare it with some of the alternatives, in particular with Bootstrap, $EVT$ and $GARCH$ methods in a similar manner than in Chapter 2. Finally, we compare some of these models when forecasting the 10-steps ahead $VaR$ using daily or fortnightly data.

This Chapter has been organized as follows. Section 3.2 present the analysis of the effects that produce in the results measuring the $VaR$ and the $ES$ with different confidence levels instead of the required 1%. Section 3.3 describe some of the methods used for estimating the $VaR$ on alternative horizons and illustrates the estimation by implementing them to estimate the $VaR$ of a series of fortnightly and daily $S\&P500$ index returns. Finally, Section 3.4 concludes the Chapter with the main conclusions and suggestions for further research.

## 3.2 Effects of different VaR and ES levels

Using the data described in Chapter 2, we follow the same procedure, first we backtest the models in order to obtain which one is the most accurate using the tests of Christoffersen (1998) in (2.19), the $DQ$ test of Engle and Manganelli (2004) and the $LB$ test of Berkowitz et al. (2006) in (2.20) computed for $m = 5$ and 20. This results are presented on Table 3.1 and 3.2 for $\alpha = 0.05$ and $\alpha = 0.1$ respectively. We observe that again the method which discriminate better among models is the test of Christoffersen (1998). Fixing the significance level to the 10%, Historical Simulation and both forms of the $CAViaR$ model are clearly rejected by all the tests for both confidence levels except for the $Asymmetric$ representation of the $CAViaR$ which is not rejected at the 5% using the $LB$ test. With respect to the rest of the models, the test of Christoffersen (1998) reject most of them at the 5% except for the $GJR$ with all the distributions except the $Skewed - t$ and the $EGARCH$ with the $Hill$ estimator. With respect to the 10% the test do not reject the models when the error distribution are assumed to be $Normal$, $Student - \nu$, $GED$ and $Skewed - t$.

Once we obtain the models which provide the most accurate forecast, we implement the procedures described in Chapter 2 for choosing the best among them. Table 3.3 shows the results of the procedure proposed by Lopez (1999) with the $RLF$, the $QLF$ and the $PQLF$ loss functions. For the 5%, using the $QLF$ and the $PQLF$, the model chosen is the $EGARCH{-}Hill$ and with the $RLF$ is the $GJR{-}Bootstrap$. On the other hand, for the 10%, with the $RLF$ we choose the $APARCH$ with $Skewed{-}t$ errors, with the $QLF$ the $EGARCH{-}t$ and with the $PQLF$ the $APARCH{-}t$. As we notice before, the values of the three loss functions are very similar for the 1% and also for the 5%.

However, for the 10% we observe that the values change a bit more, this would imply that the procedure of Lopez (1999) discriminate better for higher confidence levels. Therefore, for regulatory purposes we should use another procedure because the central interest is in the 1% level $VaR$. Alternatively, Table 3.3 also reports the p-values of the $SPA$ test in (2.26) using the three loss functions. The models which are not rejected when using any of the three loss functions are, for the 5%, the $GJR - Bootstrap$ and the $EGARCH - Hill$ and for the 10% the $GARCH - Skewed - t$, $TGARCH - GED$, $APARCH - t$ and the $APARCH - Skewed - t$. We can conclude that regardless the confidence level, the models chosen as the best are those with asymmetry involved in the specification of the conditional variance, we could expect that because of the known behavior of financial series of returns. With respect to the distribution, it seems that the assumption of asymmetry in the distribution of the residuals as well as Bootstrap and $EVT$ with the $Hill$ estimator provide better $VaR$ forecasts.

With the models chosen as the best with the $SPA$ test, we implement the test proposed by Sarma et al. (2003) to compare the models pairwise with the three loss functions. Table 3.4 reports the results for the 5% and we can observe that only when using the $QLF$ the $GJR - Bootstrap$ is rejected as inferior when compared with the $EGARCH - Hill$. Finally, Table 3.4 shows that for the 10% the $RLF$ and the $QLF$, the $TGARCH - GED$ is superior to the $APARCH - Skewed - t$. Using the $QLF$ and the $PQLF$ the $APARCH - t$ is rejected as inferior compared with the $APARCH - Skewed - t$ and with the $PQLF$ the $GARCH - Skewed - t$ and the $TGARCH - GED$ are rejected compared with the $APARCH - t$. As we mention before, this test of is not very informative, but one conclusion that we can obtain is that it seems that among the conditional volatility models, the $TGARCH$ and the $GJR$ are preferred.

With respect to the $ES$, Table 3.5 reports the p-values of the $SPA$ test with the loss function defined in equation (2.31) . For the 5% confidence level, the $GJR - Bootstrap$, $GJR - GPD$ and $EGARCH - Hill$ are not rejected to be inferior than the alternatives. In the case of the 10%, the $GARCH - N$, $TGARCH - GED$, $GJR - Bootstrap$, $EGARCH - N$, $EGARCH - Skewed - t$, $APARCH - GED$ and $APARCH - Skewed -$

$t$ are not rejected. Finally, Table 3.6 reports the results of the test proposed by Sarma et al. (2003). For the 5% none of the models are pairwise rejected. On the other hand, for the 10% confidence level, the $GARCH-N$ and the $TGARCH-GED$ are rejected compared with $GJR-Bootstrap$, and the $GJR-Bootstrap$ compared with the $EGARCH-N$, $EGARCH-Skewed-t$, $APARCH-GED$ and $APARCH-Skewed-t$. Again we notice that in the case of the $ES$ the preferred models are those with asymmetry in the conditional volatility and in the distribution, although in this case, the $GED$ distribution is preferred as well.

Now, we focus our attention in one of the principal conclusions of Chapter 2, and we want to assess if the results still valid regardless the confidence level. Figure 3.1 and Figure 3.8 represent scatter-plots of the VaR estimated with a Normal distribution and the models considered for the conditional volatility for the 5% and 10% confidence level respectively. We can observe the same results than in the 1% case, the estimations obtained with the $TGARCH$, $GJR$ and $APARCH$ models are almost the same. However, the $GARCH$ model provides different estimations. Similarly, Figure 3.2 and Figure 3.9 show the $VaR$ estimated using bootstrap. In this case, we obtain the same conclusion than before, the $GARCH$ is the only model with clearly different $VaR$ estimations.

Figure 3.3 and Figure 3.10 show scatterplots of the $VaR$ estimated using the $EGARCH$ model and all the alternative distributions. We observe similar results than those obtained in the 1% case, the $VaR$ estimated at the 5% and 10% are very similar regardless the assumed distribution. However, the differences are larger when the distribution is fixed and the model changes.

Additionally, Figure 3.4 and Figure 3.11 compare the behavior of the Asymmetric $CAViaR$ and the $EGARCH$ model with all distributions of the errors. We observe that similar to the 1% case, the Asymmetric $CAViaR$ estimated at the 5% and 10% confidence level is less conservative than the $EGARCH$ model because is greater in absolute value than the $VaR$ calculated with other models.

Finally, Figure 3.5 and Figure 3.12 show the estimates obtained with Historical Simulation, both assumptions of the $CAViaR$ method and the $EGARCH$ with all distributions considered. As well as with the 1% confidence level, $HS$ and Adaptive $CAViaR$ are almost constant over time and as we conclude in Figure 3.4 and Figure 3.11, the Asymmetric $CAViaR$ has lower estimations than the $EGARCH$.

Following the same analysis, Figure 3.6 and Figure 3.13 represent scatter plots of the estimations of the $ES$ assuming Normal errors and the five models for the conditional volatility. Again, we observe that only the estimations of the $GARCH$ model are different from those obtained with the rest of the models. Figure 3.7 and Figure 3.14 show scatterplots of the estimations of the $ES$ for the $EGARCH$ model using all distributions of the errors. We observe that the estimations are very similar for all distributions.

Therefore, we can conclude that regardless the confidence level and at least for the $S\&P500$ returns, the adequate specification of the variance is more important than the assumption for the error distribution for both measures of risk.

## 3.3 Alternative VaR and ES horizons

This section describes some of the existing alternative methods for estimating long horizon $VaR$ and $ES$. We focus on the ten-steps ahead forecast required by the Basel Committee. The simplest and most widely used approach is the square root of time given by (1.6).

Alternatively, we can write the ten-steps ahead estimation of the $VaR$ and $ES$ as follows

$$\widehat{VaR}^{\alpha}_{t+10} = \widehat{q}_{\alpha}\widehat{\sigma}_{t+10} \tag{3.1}$$

$$\widehat{ES}^{\alpha}_{t+10} = \widehat{\sigma}_{t+10} \underset{t-1}{E} \left[\widehat{\epsilon}_{t+10} | \widehat{\epsilon}_{t+10} \leq \widehat{q}_{\alpha}\right] \tag{3.2}$$

There are different alternative specifications for the conditional variance, for example, the $GARCH$, $TGARCH$ and $GJR$ models. Forecasts of these $GARCH$-type models are not straightforward except for the $GARCH(1,1)$ in (2.4) which has a closed form for the ten-steps ahead forecast given by

$$\widehat{\sigma}_{t+10} = \frac{\alpha_0 \left[1 - (\alpha_1 + \beta_1)^{10-1}\right]}{1 - \alpha_1 - \beta_1} + (\alpha_1 + \beta_1)^{10-1} \widehat{\sigma}_{t+1}.$$

In the case of the $TGARCH$ and the $GJR$ models in (2.5), we can easily obtain the one-step ahead forecast of the conditional volatility. However, for a longer horizon the volatility is unknown and we have to simulate it. One way to achieve that is simulating pathways of the ten-steps ahead distribution of the returns as done by Barone-Adesi et al. (1999). Then, this is introduced in the formulation of the conditional volatility and finally, the forecasts are obtained recursively.

On the other hand, the ten-steps ahead distribution of the standardized residuals can be assumed to be a known distribution, for example, the Normal, Student-$\nu$, GED and the Skewed-t. Alternatively, Danielsson and de Vries (2000) propose a method for estimating the quantile and the expectation based on $EVT$ assuming that the distribution of the tails is well described by equation (2.10). Analogous to equations (2.13) and (2.30), the ten-steps ahead forecasts of the $VaR$ and $ES$ are calculated as the "$1/\xi$ root of time" given by

$$\widehat{q}_\alpha = -\,(10)^{\widehat{\xi}^{(H)}} \widehat{\epsilon}_{(k+1)} \left(\frac{\alpha}{k/T}\right)^{-\widehat{\xi}^{(H)}}. \tag{3.3}$$

$$\underset{t-1}{E}\left[\widehat{\epsilon}_{t+10} | \widehat{\epsilon}_{t+10} \leq \widehat{q}_\alpha\right] = (10)^{\widehat{\xi}^{(H)}} \frac{\widehat{q}_\alpha}{1 - \widehat{\xi}^{(H)}}. \tag{3.4}$$

Inside the $EVT$ framework, McNeil and Frey (2000) propose a method for estimating the ten-steps ahead distribution of the standardized residuals based on bootstrap and $GPD$ simulations. Once the distribution is obtained, the procedure described in Chapter 2 is applied to this data; see McNeil and Frey (2000) for more details on the algorithm.

On the other hand, bootstrap methods are also used in order to forecast the ten-steps ahead distribution. Giannopoulos (2003) propose to use $FHS$ in order to obtain an estimation of $\widehat{q}_\alpha$ as the quantile of the empirical distribution of a set of standardized residual, $\left\{\epsilon^*_{t+10}\right\}$ and the expectation by calculating the average of the returns that exceed $\widehat{q}_\alpha$. Alternatively, Ruiz and Pascual (2002b) propose to use bootstrap estimations that incorporate the parameter uncertainty.

These procedures were implemented to the same data used in Chapter 2 for the empirical application. We must remark that using daily data we can not use the back-testing procedures described in Chapter 2 for selecting the most accurate models when estimating the $VaR$. The reason is because the failure process implied in those methods must be iid. Therefore, we can not compare $VaR$ predictions with overlapping ten day returns. On the other hand, we can implement those methods when using fortnightly data. The problem in this case is the lack of available data.

Table 3.7 reports the expected number of failures which, in this case, is 10 because we are backtesting the 1% $VaR$ with 1000 observations. Table 3.7 also reports the nonrejection regions for the number of failures, x. If x belongs to this interval, then the model correctly measure the $VaR$. The methods that are not rejected are the $GARCH$ model with Normal, Student-$\nu$ and Skewed-t errors, the $TGARCH$ assuming Normality and the $GPD$ using $EVT$. In order to choose the best method among them, Table 3.8 reports the results obtained by implementing the procedure proposed by Lopez (1999) and the SPA test using the $PQLF$ loss function defined in (2.23). We observe that the $GARCH-N$ is the best method using the procedure of Lopez (1999). On the other hand, with the SPA test, none of the alternatives are rejected as inferior except for the $TGARCH-GPD$. Comparing this results with those obtained for the one-step ahead $VaR$ we observe that we obtain the same results. In Chapter 2 when implementing the procedure by Lopez (1999) we chose also the $GARCH-N$ as the best model. Furthermore, with the SPA test the $GARCH-N$ and the $TGARCH-N$ were selected for estimating the one-step ahead $VaR$ as well as for the ten-steps ahead. Table 3.8 also shows that, according to the $ES$, none of the methods are rejected as inferior. This was also the case for the one-step ahead $ES$.

One problem that we face when trying to compare the procedures described above with the square root of time was that the estimations obtained with the latter were too large in absolute value. Consequently, we could not construct the failure process needed for backtesting as none of the returns were greater than these estimations. However, we made a graphical example in order to show the differences among methods. Figure 3.15 shows the estimations for the ten-steps ahead VaR and ES obtained assuming Normality with all the alternative models for the conditional volatility and those obtained using the square root of time. We clearly observe that, for all cases, the square root provides greater estimations in absolute value than the others, which illustrate the problem of overestimation mentioned in the literature by Kupiec (1995) and Blake et al. (2000).

On the other hand, we also want to asses if the results are different when using daily or fortnightly data. We implemented the previous procedures to the $S\&P500$ returns observed from 11/01/1985 to 18/04/2008. Resulting in 4421 daily observations and 421 fortnightly observations for estimation and for backtesting, 1438 daily observations and 150 fortnightly observations.

Table 3.9 reports the expected number of failures and the nonrejection regions for the number of failures in each case. Additionally, for fortnightly data, Table 3.9 shows the p-values of the test of Christoffersen (1998). In the top of this table we can observe that when implemented the procedures to daily data, the most accurate models are the $GARCH$ and $GJR$ with $Normal$, $Skewed-t$ and *Bootstrap* errors On the other hand, the bottom of Table 3.9 we observe that, with both methods, the only model rejected as accurate is the $GARCH$ with *Bootstrap* errors. Table 3.10 shows that using the procedure by Lopez (1999) the best model is the $GJR - N$ for daily data. However, when using fortnightly data the chosen model is the $GARCH - N$. We observe also differences when using the SPA test, because in the daily data case, we can not reject as inferior the $GARCH$ and $GJR$ assuming Normality and asymmetry in the distribution by the$Skewed - t$. On the other hand, for fortnightly data, the chosen models are the $GARCH$ model assuming normality and the $GPD$ for the tail of the distribution, the $TGARCH$ model with all the specifications for the distribution of the standardized residuals and the $GJR$ with $Normal$ errors.

Therefore, we can conclude that, at least for the $S\&P500$ returns, the results will change depending on the periodicity of the data we use for estimating the ten-steps ahead $VaR$.

## 3.4 Conclusions

In this Chapter we analyze if the conclusions about the chosen models when estimating the $VaR$ and $ES$ change with the confidence level and horizon. By implementing the procedures described in Chapter 2 to the same series of returns we conclude that the asymmetry involved in the specification of the conditional variance as well as the asymmetry in the distribution of the residuals, Bootstrap and $EVT$ with the $Hill$ estimator provide better $VaR$ forecasts regardless the confidence level. Furthermore, at least for the $S\&P500$ returns, the adequate specification of the variance is more important than the assumption for the error distribution for both measures of risk when assuming different confidence levels.

On the other hand, with respect to the horizon, the results obtained for the one-step ahead $VaR$ and $ES$ can be generalized to the ten-steps ahead. Furthermore, we graphically show that the square root of time overestimate both measures of risk. The use of the square root of time by financial institutions will lead them to save more capital than needed which can be reflected as lack of liquidity to face additional obligations.

Finally, we conclude that there are differences when estimating the ten-steps ahead $VaR$ using daily or fortnightly data. However, one issue related to choosing fortnightly data is the availability of information, because, according to the Basel Committee, at least 250 observations must be used for estimation, which means to have at least seven years. The same problem has to be faced for backtesting purposes, because some times it is needed to backtest over a large number of observations.

Table 3.1. p-values of Backtesting $VaR5\%$ tests for S&P500 based on n=1000 days.

| | | Christoffersen Likelihood | DQ Test | LB(5) | LB(20) |
|---|---|---|---|---|---|
| | | p-value | p-value | p-value | p-value |
| Historical Simulation | | 1.82E-05 | 4.20E-06 | 5.03E-14 | 3.62E-51 |
| CAViaR Adaptive | | 9.73E-09 | 1.09E-06 | 2.18E-10 | 4.32E-32 |
| CAViaR Asymmetric | | 0.007 | 0.027 | **0.785** | **0.306** |
| GARCH | Normal | 0.035 | **0.269** | **0.180** | **0.496** |
| | Student- v | 0.013 | 0.098 | 0.048 | **0.168** |
| | GED | 0.035 | **0.271** | **0.180** | **0.496** |
| | Skewed-t | 0.009 | **0.107** | 0.071 | **0.297** |
| | Bootstrap | 0.051 | **0.161** | **0.662** | **0.154** |
| | EVT-Hill | 0.035 | **0.269** | **0.180** | **0.496** |
| | EVT-GPD | 0.035 | **0.270** | **0.180** | **0.496** |
| TGARCH | Normal | 0.020 | **0.243** | **0.539** | **0.958** |
| | Student- v | 0.020 | **0.252** | **0.369** | **0.921** |
| | GED | 0.014 | **0.227** | **0.414** | **0.969** |
| | Skewed-t | 0.014 | **0.225** | **0.414** | **0.969** |
| | Bootstrap | 0.026 | **0.265** | **0.489** | **0.911** |
| | EVT-Hill | 0.042 | **0.280** | **0.397** | **0.847** |
| | EVT-GPD | 0.026 | **0.263** | **0.489** | **0.936** |
| GJR | Normal | **0.131** | **0.429** | **0.440** | **0.911** |
| | Student- v | **0.320** | **0.619** | **0.349** | **0.749** |
| | GED | **0.121** | **0.430** | **0.412** | **0.915** |
| | Skewed-t | 1.41E-06 | 0.001 | **0.859** | **0.720** |
| | Bootstrap | **0.220** | **0.635** | **0.554** | **0.947** |
| | EVT-Hill | **0.320** | **0.606** | **0.550** | **0.882** |
| | EVT-GPD | **0.229** | **0.500** | **0.366** | **0.919** |
| EGARCH | Normal | 0.026 | **0.211** | **0.489** | **0.931** |
| | Student- v | 0.033 | **0.222** | **0.438** | **0.960** |
| | GED | 0.068 | **0.566** | **0.455** | **0.768** |
| | Skewed-t | 0.010 | **0.196** | **0.451** | **0.963** |
| | Bootstrap | 0.049 | 0.024 | **0.039** | **0.303** |
| | EVT-Hill | **0.220** | **0.460** | **0.780** | **0.974** |
| | EVT-GPD | 0.042 | **0.274** | **0.594** | **0.937** |
| APARCH | Normal | 0.020 | **0.247** | **0.539** | **0.882** |
| | Student- v | 0.020 | **0.214** | **0.539** | **0.906** |
| | GED | 0.014 | **0.213** | **0.588** | **0.894** |
| | Skewed-t | 0.020 | **0.253** | **0.369** | **0.936** |
| | Bootstrap | 0.020 | **0.237** | **0.539** | **0.958** |
| | EVT-Hill | 0.050 | **0.303** | **0.369** | **0.756** |
| | EVT-GPD | 0.033 | **0.287** | **0.438** | **0.831** |

The shaded areas represent models that are not rejected when $\alpha$=0.10 under the different backtesting methods.

Table 3.2. p-values of Backtesting $VaR10\%$ tests for S&P500 based on n=1000 days.

| | | Christoffersen Likelihood | DQ Test | LB(5) | LB(20) |
|---|---|---|---|---|---|
| | | p-value | p-value | p-value | p-value |
| Historical Simulation | | 7.57E-05 | 1.27E-08 | 6.18E-15 | 9.12E-63 |
| CAViaR Adaptive | | 1.97E-14 | 3.76E-10 | 1.68E-09 | 5.50E-29 |
| CAViaR Asymmetric | | 1.97E-14 | 3.48E-13 | 1.68E-09 | 5.50E-29 |
| GARCH | Normal | 0.141 | 0.561 | 0.691 | 0.325 |
| | Student- v | 0.715 | 0.961 | 0.878 | 0.306 |
| | GED | 0.618 | 0.883 | 0.757 | 0.475 |
| | Skewed-t | 0.552 | 0.817 | 0.696 | 0.413 |
| | Bootstrap | 4.30E-10 | 5.90E-10 | 0.638 | 0.085 |
| | EVT-Hill | 0.005 | 0.146 | 0.703 | 0.103 |
| | EVT-GPD | 0.044 | 0.323 | 0.564 | 0.111 |
| TGARCH | Normal | 0.036 | 0.161 | 0.536 | 0.111 |
| | Student- v | 0.299 | 0.335 | 0.470 | 0.117 |
| | GED | 0.121 | 0.261 | 0.555 | 0.223 |
| | Skewed-t | 0.094 | 0.194 | 0.548 | 0.195 |
| | Bootstrap | 0.038 | 0.254 | 0.717 | 0.236 |
| | EVT-Hill | 0.005 | 0.063 | 0.750 | 0.708 |
| | EVT-GPD | 0.094 | 0.227 | 0.657 | 0.088 |
| GJR | Normal | 0.090 | 0.280 | 0.344 | 0.287 |
| | Student- v | 0.403 | 0.731 | 0.561 | 0.096 |
| | GED | 0.140 | 0.495 | 0.487 | 0.403 |
| | Skewed-t | 1.20E-06 | 2.99E-04 | 0.065 | 0.007 |
| | Bootstrap | 0.289 | 0.579 | 0.811 | 0.137 |
| | EVT-Hill | 0.003 | 0.069 | 0.945 | 0.496 |
| | EVT-GPD | 0.026 | 0.187 | 0.821 | 0.500 |
| EGARCH | Normal | 0.121 | 0.137 | 0.580 | 0.133 |
| | Student- v | 0.490 | 0.287 | 0.768 | 0.517 |
| | GED | 0.235 | 0.296 | 0.301 | 0.091 |
| | Skewed-t | 0.245 | 0.196 | 0.538 | 0.113 |
| | Bootstrap | 0.019 | 0.009 | 0.102 | 0.035 |
| | EVT-Hill | 0.019 | 0.101 | 0.469 | 0.413 |
| | EVT-GPD | 0.092 | 0.131 | 0.577 | 0.203 |
| APARCH | Normal | 0.094 | 0.273 | 0.693 | 0.398 |
| | Student- v | 0.299 | 0.242 | 0.514 | 0.250 |
| | GED | 0.156 | 0.289 | 0.578 | 0.147 |
| | Skewed-t | 0.156 | 0.301 | 0.578 | 0.217 |
| | Bootstrap | 0.027 | 0.145 | 0.694 | 0.187 |
| | EVT-Hill | 0.006 | 0.061 | 0.789 | 0.608 |
| | EVT-GPD | 0.052 | 0.163 | 0.739 | 0.277 |

The shaded areas represent models that are not rejected when $\alpha$=0.10 under the different backtesting methods.

Table 3.3. Test proposed by Lopez (1999) and SPA test to compare accurate $VaR5\%$ and $VaR10\%$ estimates computed with alternative loss functions.

| $VaR5\%$ | Loss Function | | | SPA test | | |
|---|---|---|---|---|---|---|
| Model | RLF | QLF | PQL | RLF | QLF | PQLF |
| | | | | p-value | p-value | p-value |
| $GJR-N$ | 2.4237 | 0.0081 | 0.4458 | 0.2560 | 0.0410 | 0.1260 |
| $GJR-t$ | 2.4567 | 0.0079 | 0.4431 | 0.0570 | 0.1690 | 0.8430 |
| $GJR-GED$ | 2.3425 | 0.0082 | 0.4453 | 0.9440 | 0.3430 | 0.0480 |
| $GJR-Bootstrap$ | 2.3421 | 0.0085 | 0.4454 | 0.9740 | 0.1220 | 0.4950 |
| $GJR-Hill$ | 2.4734 | 0.0082 | 0.4448 | 0.0030 | 0.0520 | 0.4320 |
| $GJR-GPD$ | 2.3801 | 0.0083 | 0.4455 | 0.5940 | 0.0380 | 0.2140 |
| $EGARCH-Hill$ | 2.4848 | 0.0067 | 0.4422 | 0.7480 | 0.5060 | 0.9860 |

| $VaR10\%$ | Loss Function | | | SPA test | | |
|---|---|---|---|---|---|---|
| Model | RLF | QLF | PQL | RLF | QLF | PQLF |
| | | | | p-value | p-value | p-value |
| $GARCH-N$ | 6.8791 | 0.0051 | 0.7738 | 0.6940 | 0.2030 | 0.0180 |
| $GARCH-t$ | 8.0973 | 0.0047 | 0.7711 | 0.0070 | 0.8230 | 0.2390 |
| $GARCH-GED$ | 7.7002 | 0.0048 | 0.7716 | 0.0160 | 0.1680 | 0.1090 |
| $GARCH-Skewed-t$ | 7.3889 | 0.0050 | 0.7702 | 0.1570 | 0.1420 | 0.1410 |
| $TGARCH-t$ | 6.3389 | 0.0049 | 0.7539 | 0.0010 | 0.2590 | 0.1280 |
| $TGARCH-GED$ | 6.1745 | 0.0050 | 0.7555 | 0.2800 | 0.3130 | 0.1710 |
| $GJR-GED$ | 6.8353 | 0.0048 | 0.7563 | 0.0010 | 0.1270 | 0.1880 |
| $GJR-Bootstrap$ | 6.1979 | 0.0051 | 0.7587 | 0.6250 | 0.1880 | 0.0260 |
| $EGARCH-N$ | 6.2418 | 0.0044 | 0.7592 | 0.8000 | 0.2300 | 0.0120 |
| $EGARCH-t$ | 6.8464 | 0.0043 | 0.7565 | 0.0020 | 0.9410 | 0.1240 |
| $EGARCH-Skewed-t$ | 6.3056 | 0.0045 | 0.7563 | 0.4000 | 0.0390 | 0.0410 |
| $APARCH-t$ | 5.9151 | 0.0049 | 0.7450 | 0.9999 | 0.2870 | 0.9600 |
| $APARCH-GED$ | 6.2185 | 0.0050 | 0.7542 | 0.1860 | 0.0900 | 0.2940 |
| $APARCH-Skewed-t$ | 5.8866 | 0.0051 | 0.7547 | 0.9930 | 0.1160 | 0.2500 |

Table 3.4. Test proposed by Sarma et al. (2003) to compare accurate estimates of $VaR5\%$ and $VaR10\%$ computed with alternative loss functions.

| $VaR$ Regulatory Loss Function | |
|---|---|
| Model | $EGARCH-Hill$ |
| $GJR-Bootstrap$ | $-0.7799$ (0.4356) |

| $VaR$ Quantile Loss Function | |
|---|---|
| Model | $EGARCH-Hill$ |
| $GJR-Bootstrap$ | 2.5486 (0.0110) |

| $VaR$ Predictive Quantile Loss Function | |
|---|---|
| Model | $EGARCH-Hill$ |
| $GJR-Bootstrap$ | 1.0837 (0.2787) |

| $VaR$ Regulatory Loss Function | | | |
|---|---|---|---|
| Model | $TGARCH-GED$ | $APARCH-t$ | $APARCH-Skewed-t$ |
| $GARCH-Skewed-t$ | 1.2962 (0.1952) | 1.4973 (0.1346) | 1.5915 (0.1118) |
| $TGARCH-GED$ | | 0.8526 (0.3940) | 1.8368 (0.0665) |
| $APARCH-t$ | | | $-0.1014$ (0.9192) |

| $VaR$ Quantile Loss Function | | | |
|---|---|---|---|
| Model | $TGARCH-GED$ | $APARCH-t$ | $APARCH-Skewed-t$ |
| $GARCH-Skewed-t$ | 0.0756 (0.9397) | 0.2737 (0.7843) | $-0.3208$ (0.7484) |
| $TGARCH-GED$ | | 1.6022 (0.1094) | $-3.4027$ (0.0007) |
| $APARCH-t$ | | | $-3.7142$ (0.0002) |

| $VaR$ Predictive Quantile Loss Function | | | |
|---|---|---|---|
| Model | $TGARCH-GED$ | $APARCH-t$ | $APARCH-Skewed-t$ |
| $GARCH-Skewed-t$ | 1.4179 (0.1565) | 2.4255 0.0155 | 1.5650 (0.1179) |
| $TGARCH-GED$ | | 2.6647 (0.0078) | 0.6372 (0.5241) |
| $APARCH-t$ | | | $-2.4355$ (0.0150) |

*p-values in parentheses

Table 3.5. SPA test to compare $ES5\%$ and $ES10\%$ estimates obtained from accurate $VaR$ estimates.

| | SPA test |
|---|---|
| Model | RLF |
| | p-value |
| $GJR-N$ | 0.0010 |
| $GJR-t$ | 0.0010 |
| $GJR-GED$ | 0.0010 |
| $GJR-Bootstrap$ | 0.7800 |
| $GJR-Hill$ | 0.0160 |
| $GJR-GPD$ | 0.7820 |
| $EGARCH-Hill$ | 0.6200 |

| | SPA test |
|---|---|
| Model | RLF |
| | p-value |
| $GARCH-N$ | 0.3300 |
| $GARCH-t$ | 0.0230 |
| $GARCH-GED$ | 0.0780 |
| $GARCH-Skewed-t$ | 0.0680 |
| $TGARCH-t$ | 0.0730 |
| $TGARCH-GED$ | 0.1670 |
| $GJR-GED$ | 0.0150 |
| $GJR-Bootstrap$ | 0.5640 |
| $EGARCH-N$ | 0.2670 |
| $EGARCH-t$ | 0.0670 |
| $EGARCH-Skewed-t$ | 0.1280 |
| $APARCH-t$ | 0.0790 |
| $APARCH-GED$ | 0.1380 |
| $APARCH-Skewed-t$ | 0.1110 |

Table 3.6. Test proposed by Sarma et al. (2003) to compare accurate estimates of $VaR5\%$ and $VaR10\%$ using Angelidis and Degiannakis (2007) loss function.

| $ES5\%$ Regulatory Loss Function | | |
|---|---|---|
| Model | $GJR-GPD$ | $EGARCH-Hill$ |
| $GJR-Bootstrap$ | −0.0039 (0.9969) | −1.0311 (0.3027) |
| $GJR-GPD$ | | −1.0246 (0.3058) |

| $ES10\%$ Regulatory Loss Function | | | | | | |
|---|---|---|---|---|---|---|
| Model | $TGARCH-GED$ | $GJR-Bootstrap$ | $EGARCH-N$ | $EGARCH-Skewed-t$ | $APARCH-GED$ | $APARCH-Skewed-t$ |
| $GARCH-N$ | −0.0693 (0.9447) | 1.7747 (0.0762) | 0.0605 (0.9517) | −0.2011 (0.8406) | −0.2203 (0.8257) | −0.3107 (0.7560) |
| $TGARCH-GED$ | | 2.1681 (0.0304) | 0.2368 (0.8128) | −0.2223 (0.8241) | −0.4731 (0.6362) | −0.7686 (0.4423) |
| $GJR-Bootstrap$ | | | −2.0037 (0.0454) | −2.2977 (0.0218) | −2.2739 (0.0232) | −2.3772 (0.0176) |
| $EGARCH-N$ | | | | −0.6900 (0.4903) | −0.4783 (0.6325) | −0.5857 (0.5582) |
| $EGARCH-Skewed-t$ | | | | | −0.0178 (0.9857) | −0.1566 (0.8755) |
| $APARCH-GED$ | | | | | | −0.2454 (0.8061) |

*p-values in parentheses

Figure 3.1. Scatter-plots of the $VaR5\%$ estimated by the alternative conditionally heteroscedastic models with $Normal$ errors.

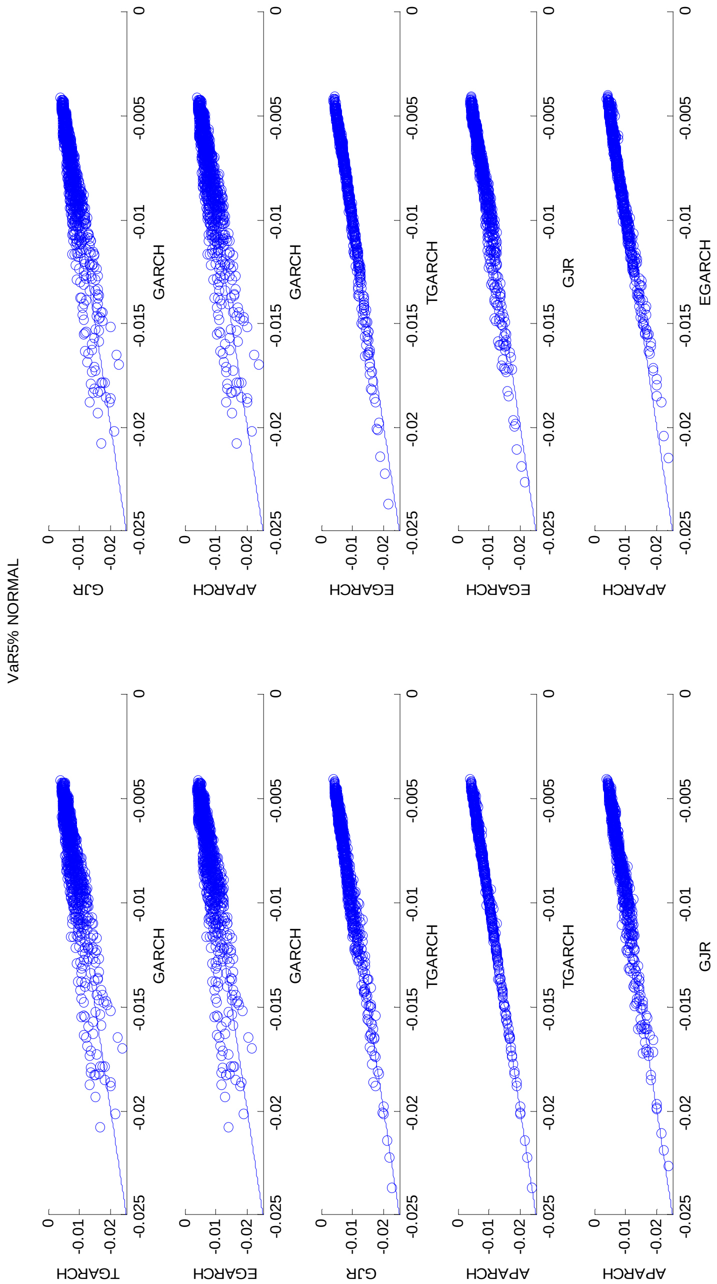


Figure 3.2. Scatter-plots of the $VaR5\%$ estimated by the alternative conditionally heteroscedastic models with *Bootstrap* errors.

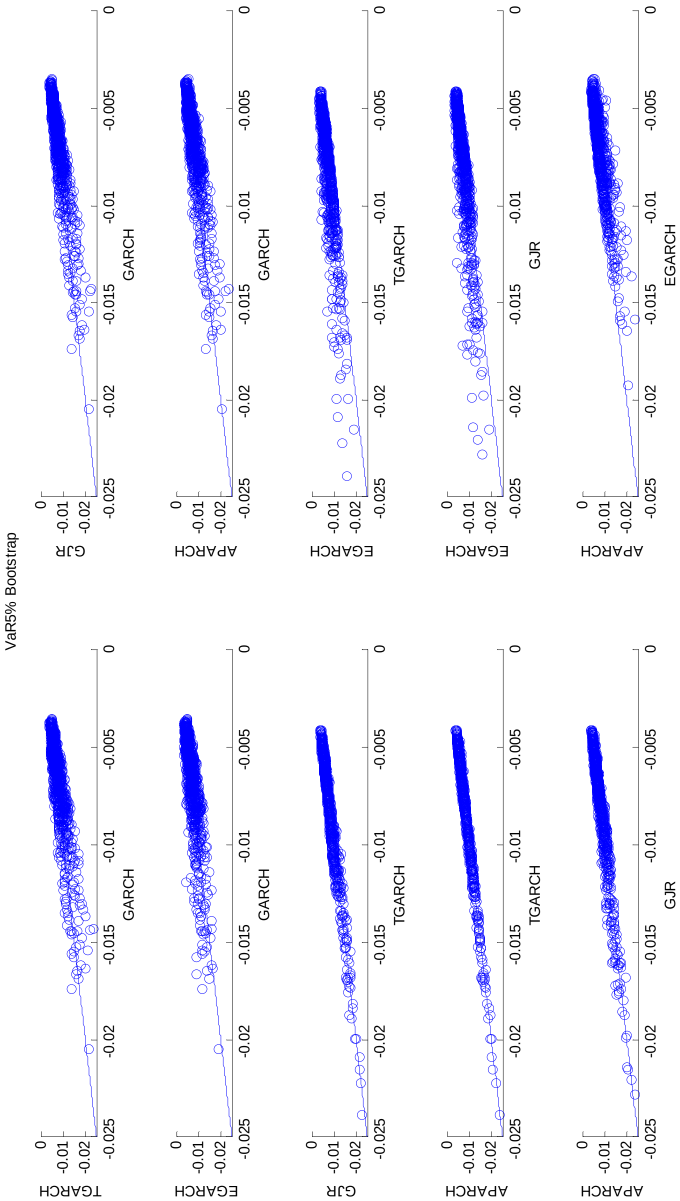


Figure 3.3. Scatter-plots of the $VaR5\%$ estimated by the alternative error distributions with the $EGARCH$ model.

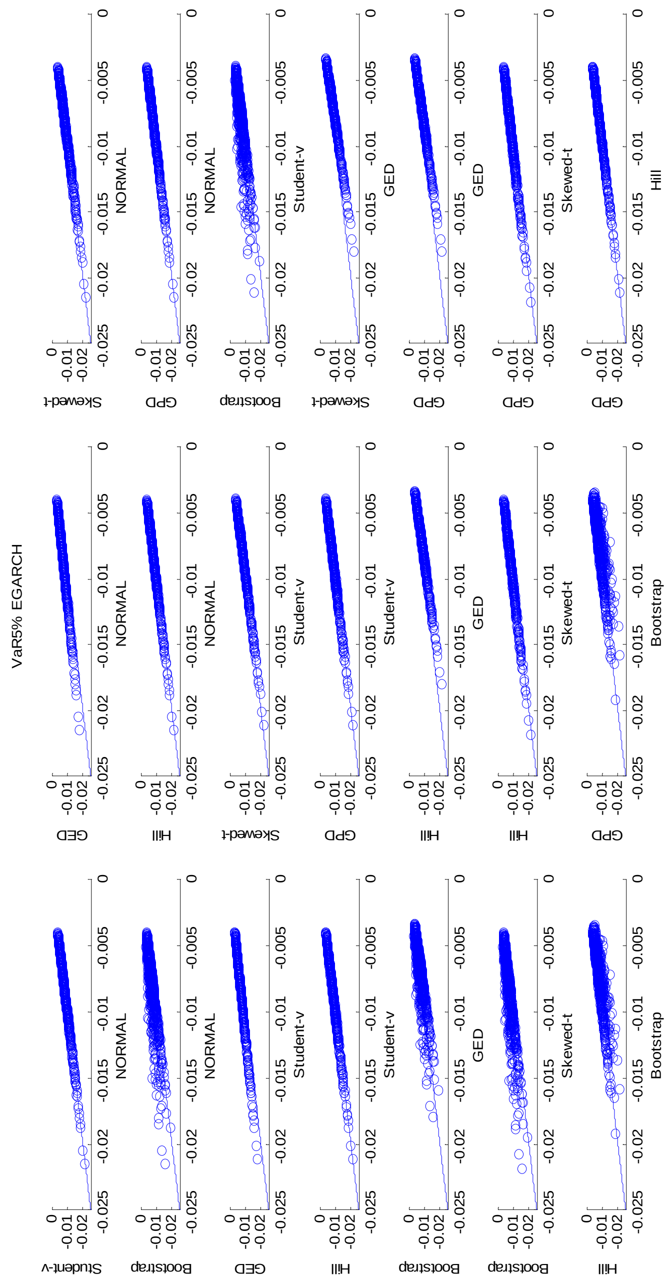


Figure 3.4. Scatter-plots of the $VaR5\%$ estimated with the $CAViaR$ $Asymmetric$ and the $EGARCH$ model with different error distributions.

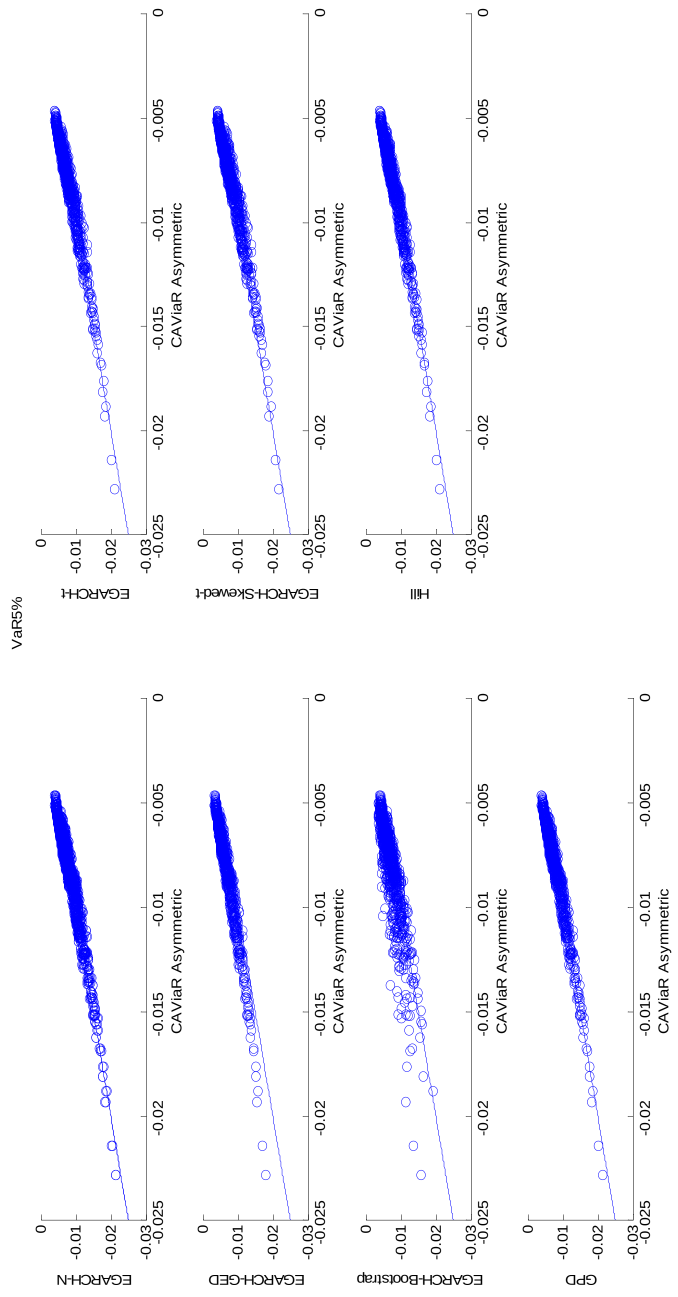


Figure 3.5. Estimates of the $VaR5\%$ obtained using the $CAViaR$ and $HS$ procedures and the $EGARCH$ model with different error distributions.

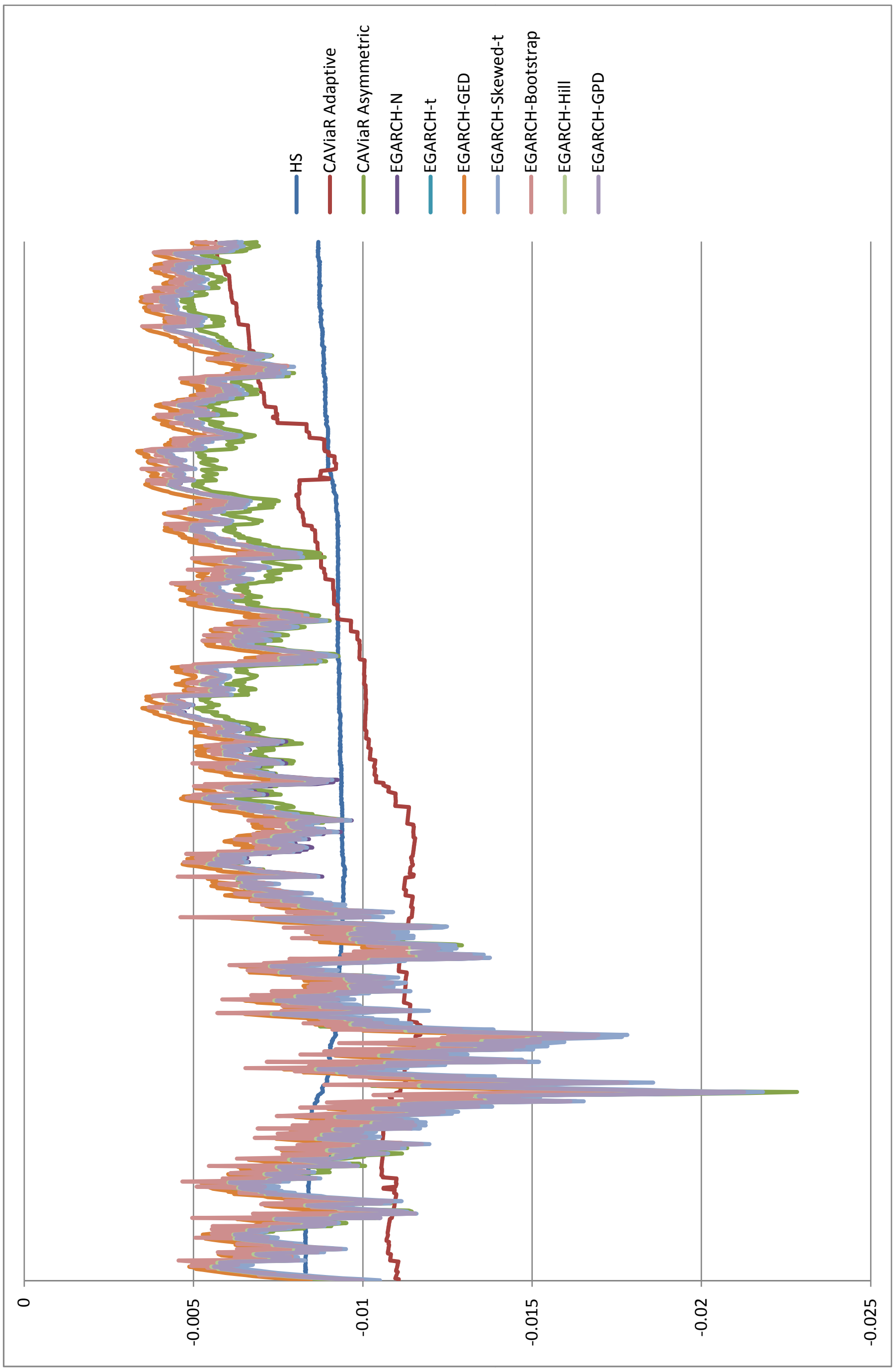


Figure 3.6. Scatter-plots of the $ES5\%$ estimated by the alternative conditionally heteroscedastic models with $Normal$ errors.

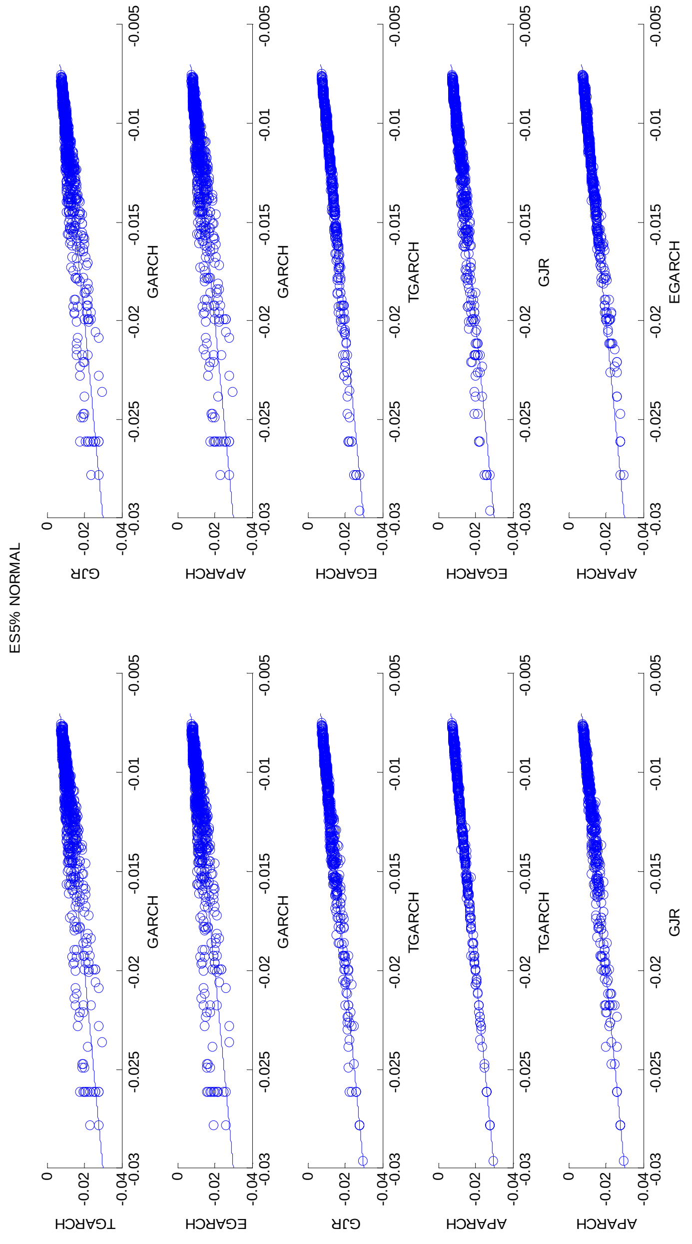


Figure 3.7. Scatter-plots of the $ES5\%$ estimated by the alternative error distributions with the $EGARCH$ model.

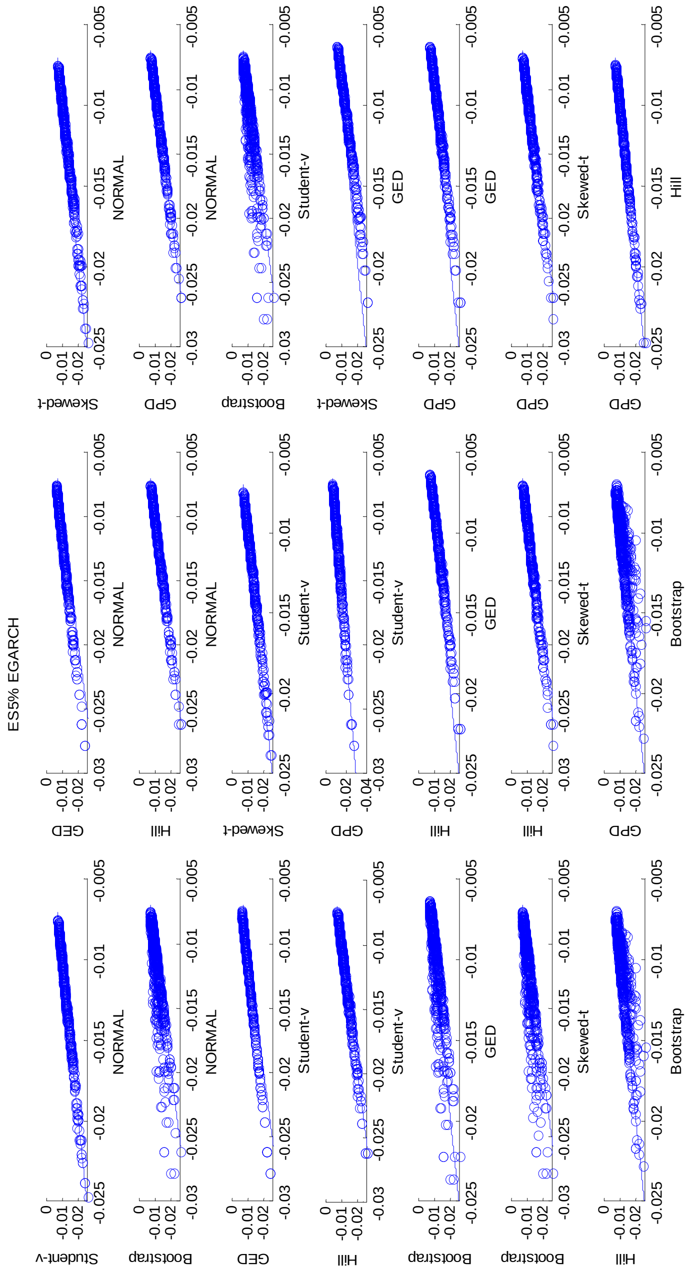


Figure 3.8. Scatter-plots of the $VaR10\%$ estimated by the alternative conditionally heteroscedastic models with $Normal$ errors.

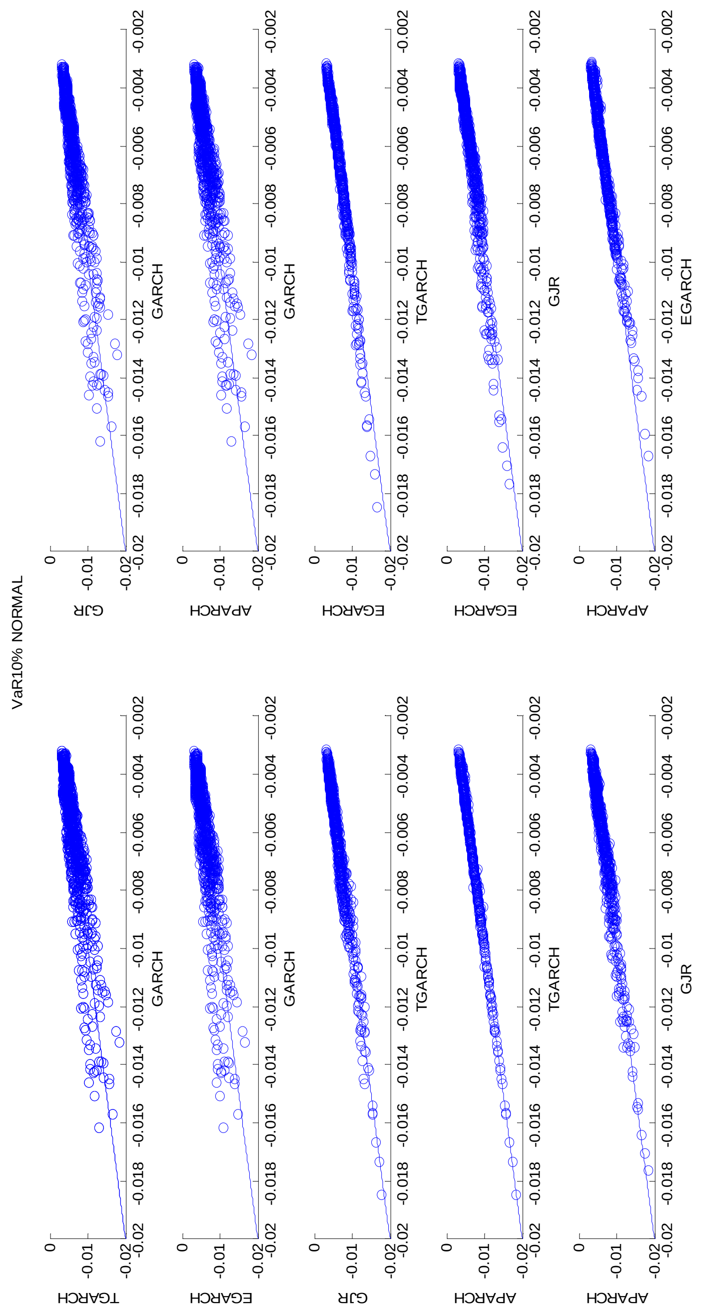


Figure 3.9. Scatter-plots of the $VaR10\%$ estimated by the alternative conditionally heteroscedastic models with *Bootstrap* errors.

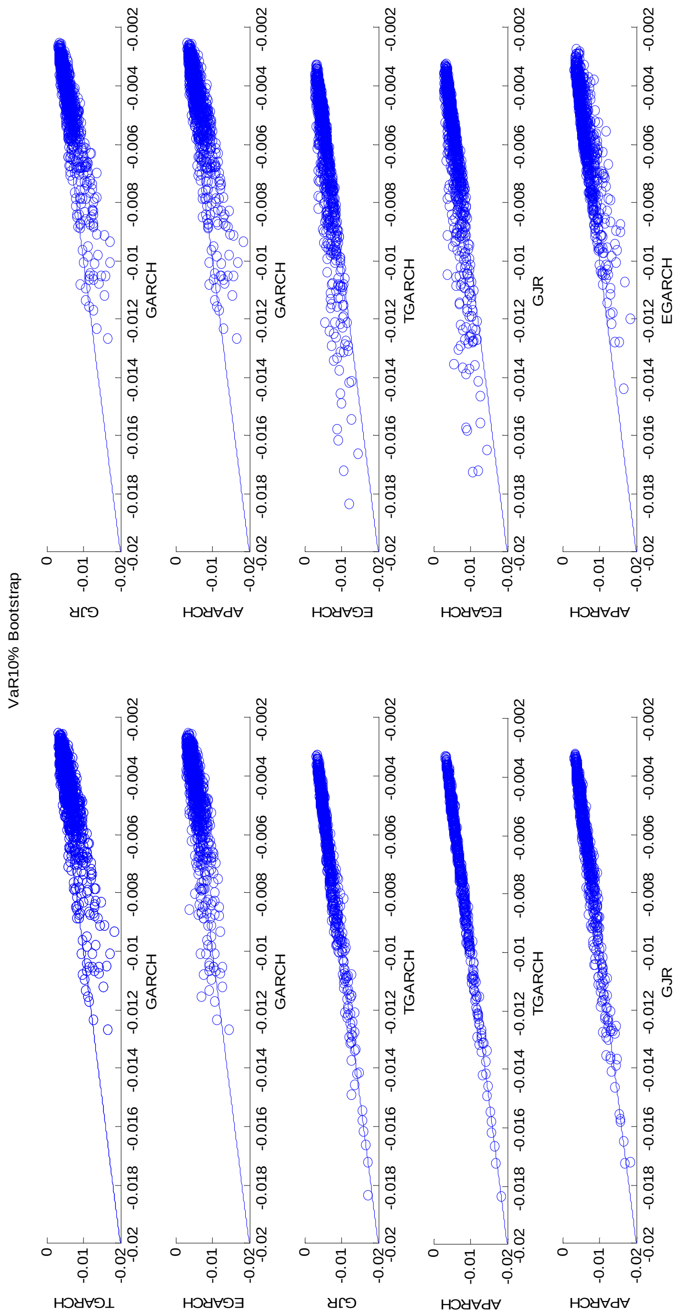


Figure 3.10. Scatter-plots of the $VaR10\%$ estimated by the alternative error distributions with the $EGARCH$ model.

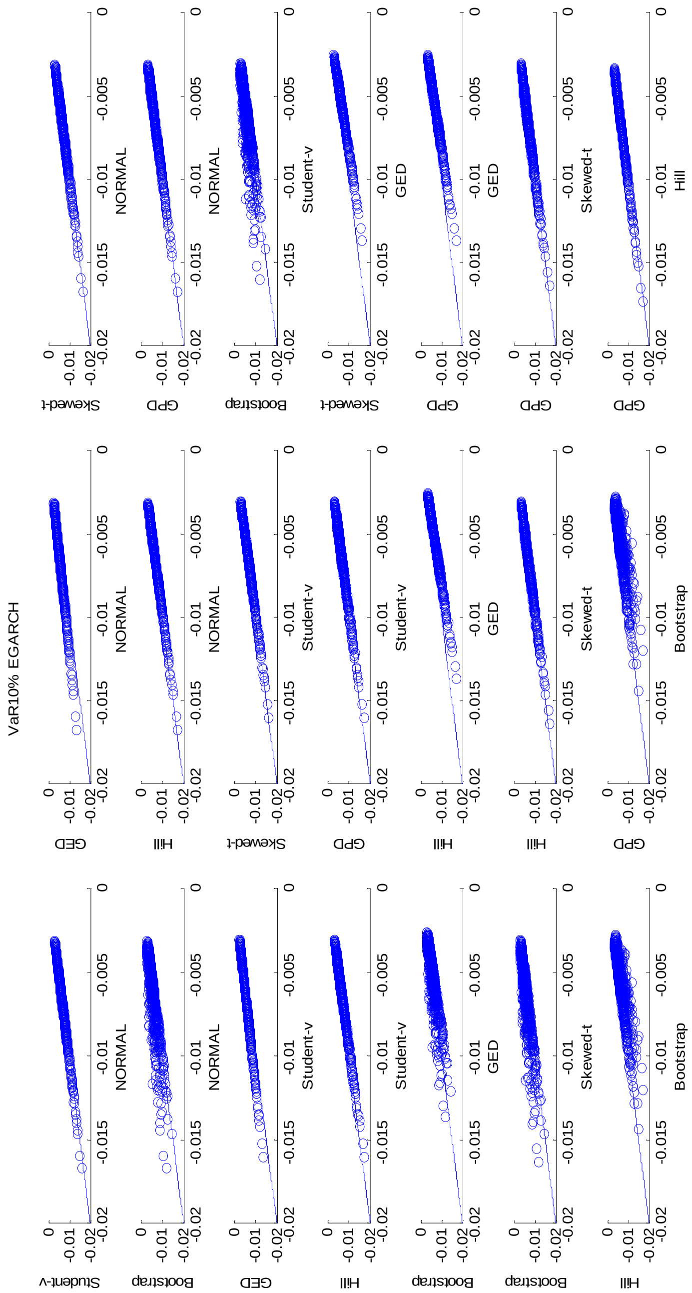


Figure 3.11. Scatter-plots of the $VaR10\%$ estimated with the $CAViaR\ Asymmetric$ and the $EGARCH$ model with different error distributions.

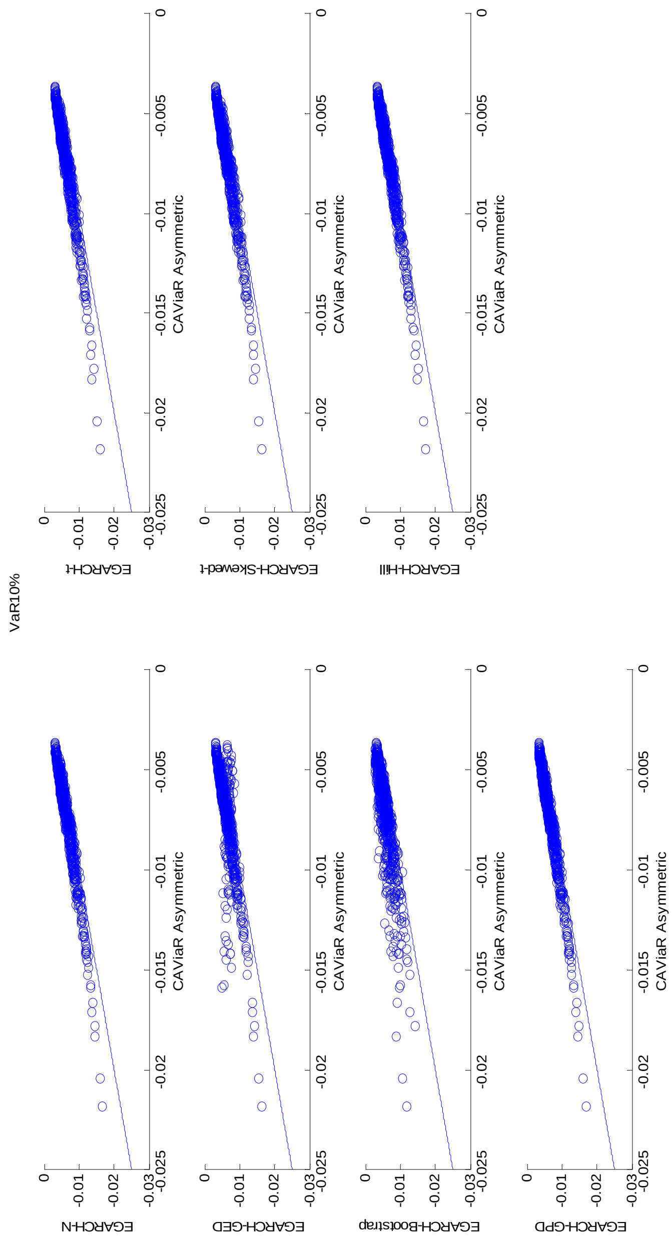


Figure 3.12. Estimates of the $VaR10\%$ obtained using the $CAViaR$ and $HS$ procedures and the $EGARCH$ model with different error distributions.

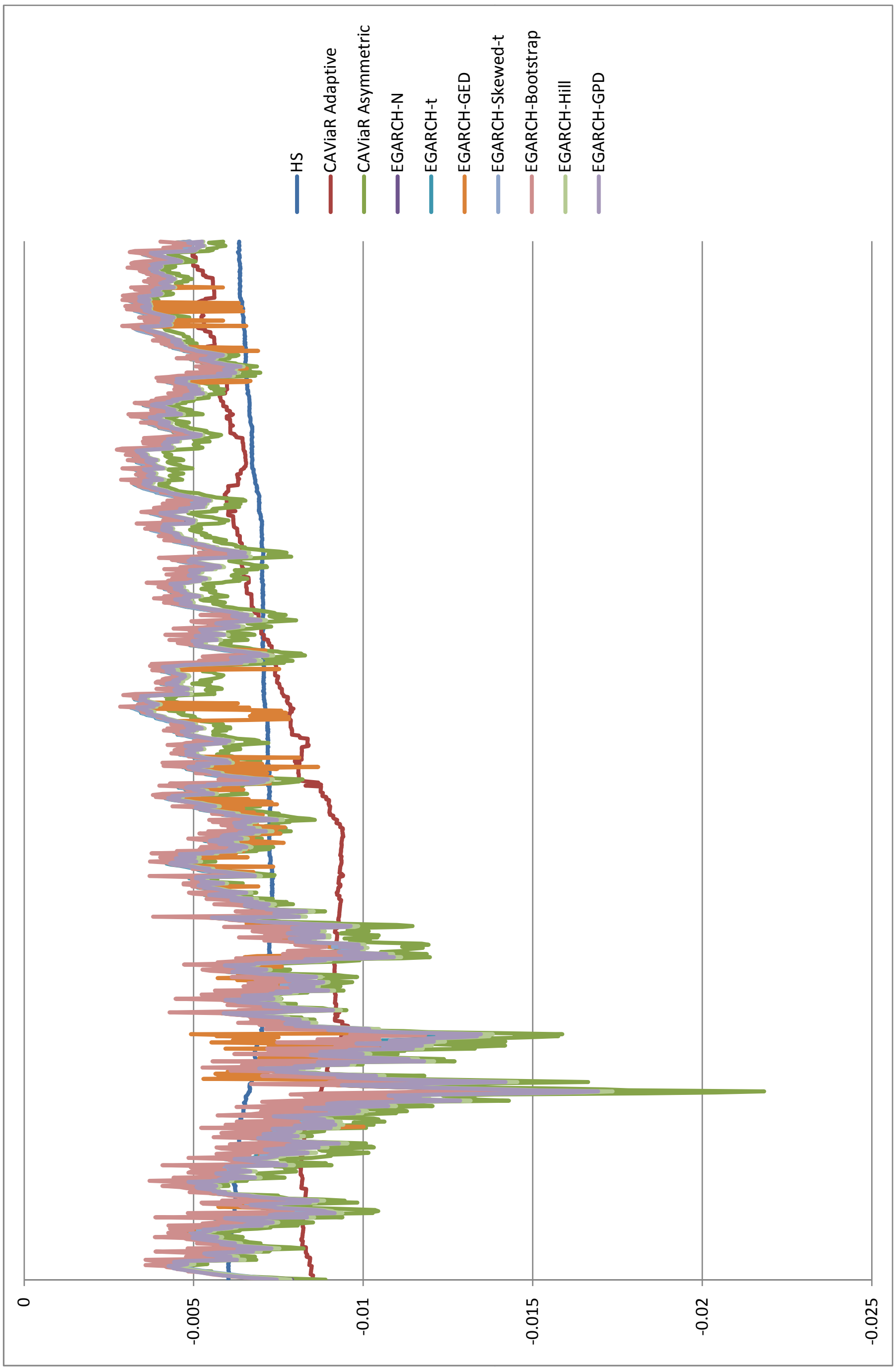


Figure 3.13. Scatter-plots of the $ES10\%$ estimated by the alternative conditionally heteroscedastic models with $Normal$ errors.

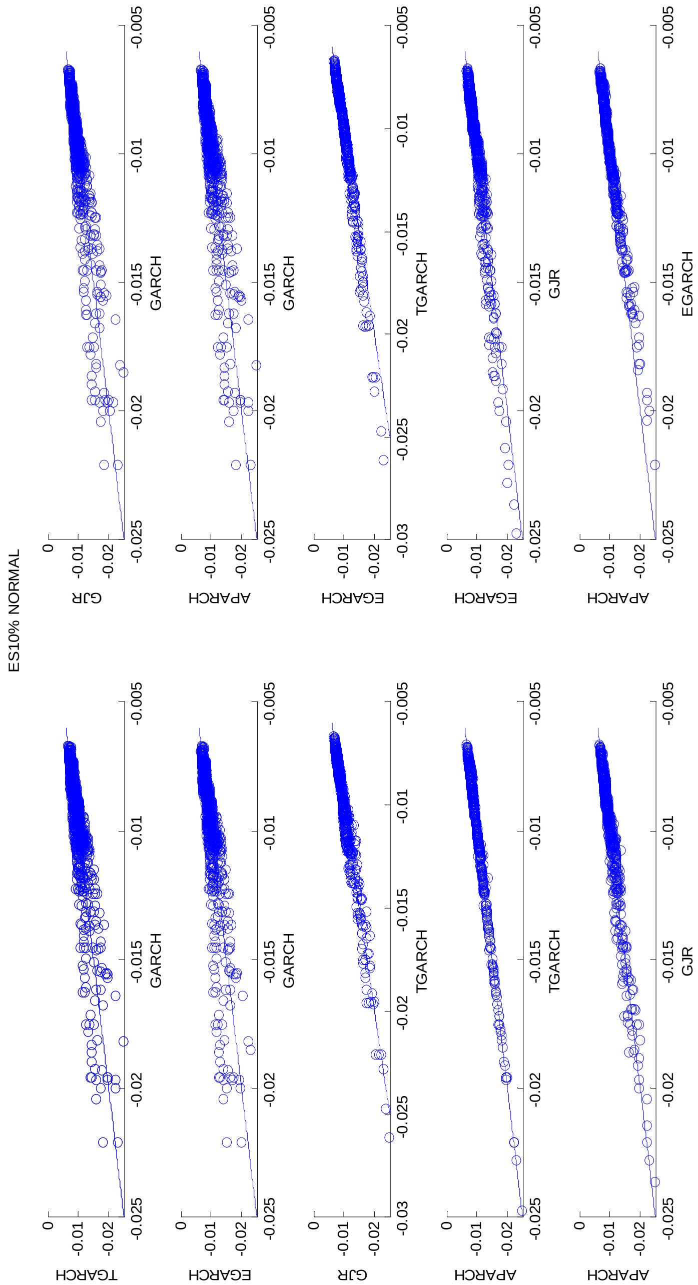


Figure 3.14. Scatter-plots of the $ES10\%$ estimated by the alternative error distributions with the $EGARCH$ model.

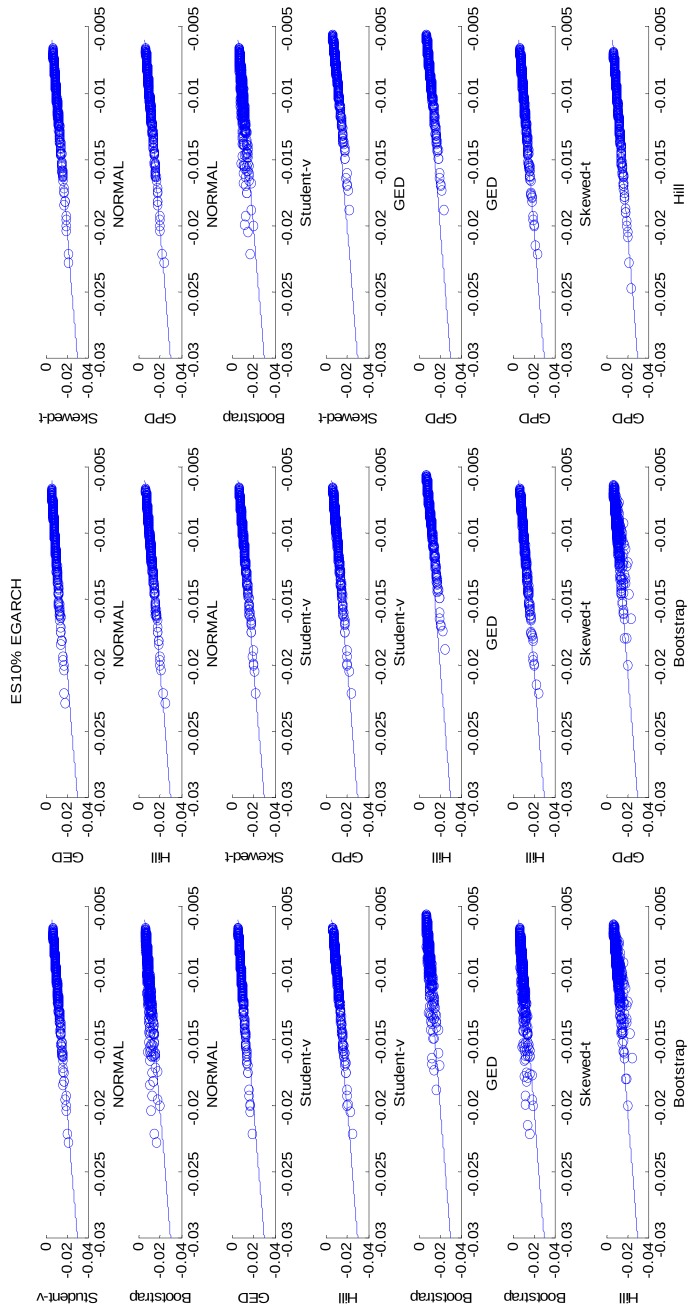


Table 3.7. Nonrejection regions for backtesting ten-steps ahead $VaR1\%$ for S&P500 based on n=1000 days.

| | | |
|---|---|---|
| Expected Number of Failures | | 10 |
| Nonrejection Region for Number of Failures x | | 3<x<16 |
| | | |
| Model | | Number of Failures |
| GARCH | Normal | 5 |
| | Student- v | 4 |
| | GED | 4 |
| | Skewed-t | 1 |
| | Bootstrap | 1 |
| | EVT-Hill | 1 |
| | EVT-GPD | 1 |
| TGARCH | Normal | 4 |
| | Student- v | 17 |
| | GED | 3 |
| | Skewed-t | 3 |
| | Bootstrap | 2 |
| | EVT-Hill | 1 |
| | EVT-GPD | 4 |
| GJR | Normal | 3 |
| | Student- v | 3 |
| | GED | 2 |
| | Skewed-t | 1 |
| | Bootstrap | 3 |
| | EVT-Hill | 1 |
| | EVT-GPD | 2 |

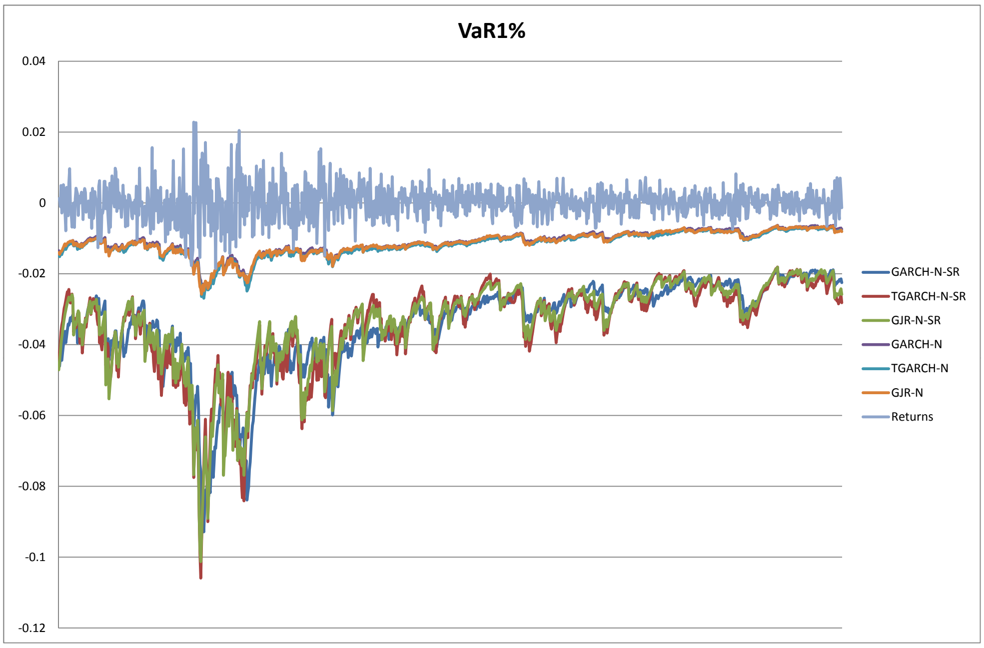
The shaded areas represent models that are not rejected when $\alpha$=0.05.

Table 3.8. Test proposed by Lopez (1999) and SPA test to compare accurate ten-steps ahead $VaR1\%$ and $ES1\%$ estimates computed with the PQLF loss function.

| $VaR1\%$ | Loss Function | SPA test |
|---|---|---|
| Model | PQLF | PQLF |
| | | p-value |
| $GARCH-N$ | 0.1191 | 0.8610 |
| $GARCH-t$ | 0.1216 | 0.1700 |
| $GARCH-GED$ | 0.1216 | 0.2000 |
| $TGARCH-N$ | 0.1216 | 0.1190 |
| $TGARCH-GPD$ | 0.1368 | 0.0010 |

| $ES1\%$ | SPA test |
|---|---|
| Model | RLF |
| | p-value |
| $GARCH-N$ | 0.9680 |
| $GARCH-t$ | 0.2110 |
| $GARCH-GED$ | 0.4400 |
| $TGARCH-N$ | 0.9310 |
| $TGARCH-GPD$ | 0.1440 |

Figure 3.15. Ten-step ahead $VaR1\%$ and $ES1\%$ estimated by the alternative conditionally heteroscedastic models with $Normal$ errors and with the square root of time.



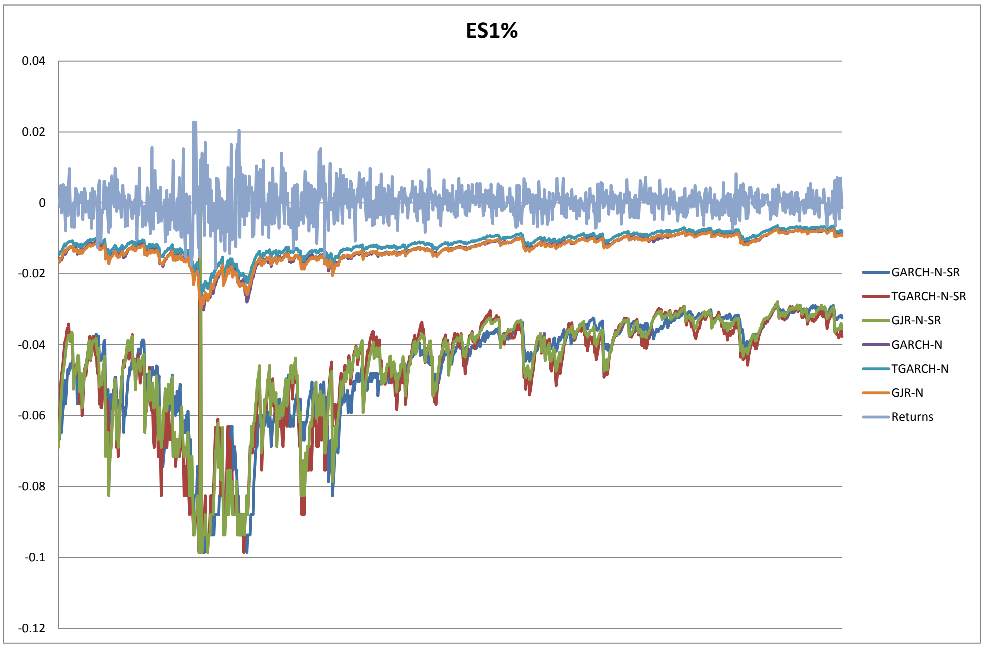



Table 3.9. Nonrejection regions and p-values of Christoffersen (1998) test for back-testing ten-steps ahead $VaR1\%$ for daily and fortnightly S&P500.

| | | |
|---|---|---|
| Expected Number of Failures | | 14 |
| Nonrejection Region for Number of Failures x | | 6<x<21 |
| Model | | Number of Failures |
| GARCH | Normal | 17 |
| | Skewed-t | 9 |
| | Bootstrap | 8 |
| | EVT-GPD | 53 |
| TGARCH | Normal | 32 |
| | Skewed-t | 41 |
| | Bootstrap | 3 |
| | EVT-GPD | 46 |
| GJR | Normal | 13 |
| | Skewed-t | 9 |
| | Bootstrap | 9 |
| | EVT-GPD | 47 |

| | | | |
|---|---|---|---|
| Expected Number of Failures | | 2 | |
| Nonrejection Region for Number of Failures x | | 0<x<3 | |
| Model | | Number of Failures | Christoffersen Likelihood p-value |
| GARCH | Normal | 2 | 0.890 |
| | Skewed-t | 1 | 0.897 |
| | Bootstrap | 9 | 7.79E-05 |
| | EVT-GPD | 2 | 0.890 |
| TGARCH | Normal | 2 | 0.890 |
| | Skewed-t | 2 | 0.890 |
| | Bootstrap | 1 | 0.890 |
| | EVT-GPD | 2 | 0.890 |
| GJR | Normal | 2 | 0.890 |
| | Skewed-t | 1 | 0.890 |
| | Bootstrap | 1 | 0.909 |
| | EVT-GPD | 2 | 0.890 |

The shaded areas represent models that are not rejected when $\alpha$=0.05.

Table 3.10. Test proposed by Lopez (1999) and SPA test to compare accurate ten-steps ahead $VaR1\%$ estimates computed with the PQLF loss function for daily and fortnightly S&P500.

| daily | Loss Function | SPA test |
|---|---|---|
| Model | PQLF | PQLF |
| | | p-value |
| $GARCH-N$ | 0.4029 | 0.3410 |
| $GARCH-Skewed-t$ | 0.4130 | 0.2160 |
| $GARCH-Bootstrap$ | 0.4294 | 0.0010 |
| $GJR-N$ | 0.3977 | 0.8970 |
| $GJR-Skewed-t$ | 0.4144 | 0.1830 |
| $GJR-Bootstrap$ | 0.4291 | 0.0010 |

| fortnightly | Loss Function | SPA test |
|---|---|---|
| Model | PQLF | PQLF |
| | | p-value |
| $GARCH-N$ | 0.1184 | 0.9650 |
| $GARCH-Skewed-t$ | 0.1308 | 0.0390 |
| $GARCH-GPD$ | 0.1216 | 0.3260 |
| $TGARCH-N$ | 0.1220 | 0.4600 |
| $TGARCH-Skewed-t$ | 0.1268 | 0.8129 |
| $TGARCH-Bootstrap$ | 0.1192 | 0.9900 |
| $TGARCH-GPD$ | 0.1276 | 0.2180 |
| $GJR-N$ | 0.1266 | 0.4080 |
| $GJR-Skewed-t$ | 0.1369 | 0.0020 |
| $GJR-Bootstrap$ | 0.2233 | 0.0010 |
| $GJR-GPD$ | 0.1329 | 0.0630 |

# Chapter 4

# Bootstrap Prediction Intervals for Risk Measures in the context of GARCH models

## 4.1 Introduction

In this Chapter, we focus on the parametric specification of returns assuming that they are represented by $GARCH$-type models and propose a new bootstrap procedure that incorporates a second bootstrap step in the estimation of the quantile of the conditional distribution of standardized returns. Using simulated and real data, we compare the performance of the new bootstrap procedure implemented when estimating the quantile by $FHS$, $EVT$ and the Gram-Charlier-Cornish-Fisher approximation with that of the bootstrap procedure proposed by Christoffersen and GonÇalves (2005). We show that, although our bootstrap procedure is very simple from a computational point of view, incorporating this second bootstrap step, improves the performance of the prediction intervals of the $VaR$ and $ES$ which have coverage much closer to the nominal. Furthermore, following Ho and Lee (2005), we also consider bootstrap prediction intervals for the quantile that overcome the limitations of the traditional prediction intervals. The iterative smoothed bootstrap of Ho and Lee (2005) may have better coverages depending on the particular value of a smoothing parameter that has to be arbitrarily chosen. However, choosing this parameter is very costly from a computational point of view increasing largely the complexity of the procedure. Furthermore, the procedure proposed by Ho and Lee (2005) can only be implemented to obtain prediction intervals for the $VaR$ but not for the $ES$.

This Chapter is organized as follows. In Section 4.2 we propose a new bootstrap procedure to obtain prediction intervals for the $VaR$ and $ES$ in the context of univariate $GARCH(1,1)$ models. Section 4.3 reports the results of several Monte Carlo experiments carried out to analyze the finite sample performance of the proposed intervals and to compare them with alternative bootstrap intervals previously proposed in the literature. Section 4.4 illustrates the proposed procedures by implementing them to obtain prediction intervals of future $VaR$ and $ES$ of several real time series of financial returns. Finally, Section 4.5 concludes the Chapter.

## 4.2 Bootstrap prediction intervals for VaR and ES

Given its simplicity and popularity in this section, we focus on the $GARCH(1,1)$ model to describe the dynamic evolution of the conditional variances of financial returns. However, the procedures described can be easily implemented for alternative specifications of the conditional variance as far as it is observable one-step ahead. Consider that the series of returns, $R_t$, is given by the following uncorrelated $GARCH(1,1)$ process,

$$R_t = \epsilon_t \sigma_t \tag{4.1}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 R_{t-1}^2 + \beta \sigma_{t-1}^2, \tag{4.2}$$

for $t = 2, ..., T$, where $\sigma_t$ is the conditional standard deviation of returns and $\sigma_1^2 = \dfrac{\alpha_0}{(1 - \alpha_1 - \beta)}$ is the marginal variance. The parameters $\alpha_0$, $\alpha_1$ and $\beta$ are assumed to satisfy the usual positivity and stationarity restrictions. The disturbances $\epsilon_t$ are assumed to be $iid$ with zero mean and variance 1. If returns are given by (4.1), the one-step ahead $VaR$ and $ES$ are given by

$$VaR_t = \sigma_t q \tag{4.3}$$

and

$$ES_t = \sigma_t \mathop{E}_{t-1} \left[ \epsilon_t \mid \epsilon_t \leq q \right], \tag{4.4}$$

respectively, where $q$ is the 1% quantile of the distribution of $\epsilon_t$. Expression (4.3), shows that in the parametric framework considered in this Chapter, we can express the $VaR_t$ as the product of the conditional standard deviation, $\sigma_t$, and a constant, $q$, which depends on the distribution of the standardized returns, $\epsilon_t$. Furthermore, the $ES$ also depends on $\sigma_t$, $q$ and on the expectation of the returns under $q$. By assuming a particular distribution of returns as, for example, the Normal or a Student-$\nu$ distribution, $q$ and the expectation involved in (4.4) have known values; see Christoffersen and GonÇalves (2005) for the corresponding expressions. However, in the general case, when the distribution of $\epsilon_t$ is unknown, $q$ and the expectation in (4.4) have to be estimated.

In any case, even if $q$ is known, one needs to estimate the parameters of the conditional variance. Due to its well known asymptotic properties, in this Chapter, we consider the Quasi Maximum Likelihood ($QML$) estimator, denoted by $\left\{\widehat{\alpha}_0, \widehat{\alpha}_1, \widehat{\beta}\right\}$. Then, in practice, the $VaR_t$ and $ES_t$ are estimated as follows

$$\widehat{VaR}_t = \widehat{\sigma}_t \widehat{q} \tag{4.5}$$

and

$$\widehat{ES}_t = \widehat{\sigma}_t \widehat{E}, \tag{4.6}$$

where $\widehat{\sigma}_t^2 = \widehat{\alpha}_0 + \widehat{\alpha}_1 R_{t-1}^2 + \widehat{\beta}\widehat{\sigma}_{t-1}^2$, for $t = 2, ..., T$ with $\widehat{\sigma}_1^2 = \dfrac{\widehat{\alpha}_0}{\left(1 - \widehat{\alpha}_1 - \widehat{\beta}\right)}$ and $\widehat{E} = \underset{t-1}{E}\left[\widehat{\epsilon_t \mid \epsilon_t} \leq q\right]$. Therefore, if the $GARCH$ model is correctly specified, there are two sources of uncertainty associated with predicting $VaR_{T+1}$ given $\{R_1, ..., R_T\}$. One is the uncertainty in computing $q$ and the other concerns the prediction of the volatility, $\sigma_{T+1}$. Furthermore, when computing the $ES$ one also need to estimate the expectation beyond $q$. In this section, we describe bootstrap procedure proposed by Christoffersen and GonÇalves (2005) to construct prediction intervals of the $VaR$ and $ES$ that take into account these two sources of uncertainty and propose an alternative procedure with better small samples properties.

### 4.2.1 Bootstrap based prediction intervals for VaR and ES

Consider the $GARCH\,(1, 1)$ model in equations (4.1) and (4.2) whose parameters have been estimated by $QML$. Then, one can obtain the standardized residuals, $\widehat{\epsilon}_t = \dfrac{R_t}{\widehat{\sigma}_t}$ where $\widehat{\sigma}_t$ is defined as in equations (4.5) and (4.6). Pascual et al. (2006) propose to obtain a bootstrap replicate of the original returns, $R_t^{*(i)}$, $i = 1, ...B$, from the following recursions:

$$\sigma_t^{*2(i)} = \widehat{\alpha}_0 + \widehat{\alpha}_1 R_{t-1}^{*2(i)} + \widehat{\beta}\sigma_{t-1}^{*2(i)} \tag{4.7}$$

$$R_t^{*(i)} = \epsilon_t^{*(i)}\sigma_t^{*(i)} \tag{4.8}$$

for $t = 2, ..., T$, where $\sigma_1^{*2(i)} = \dfrac{\widehat{\alpha}_0}{\left(1 - \widehat{\alpha}_1 - \widehat{\beta}\right)}$ and $\epsilon_t^{*(i)}$ are random draws with replacement from the standardized residuals $\widehat{\epsilon}_t$. Then, the parameters $\left(\widehat{\alpha}_0^{*(i)}, \widehat{\alpha}_1^{*(i)}, \widehat{\beta}^{*(i)}\right)$ are estimated from $\left\{R_1^{*(i)}, ..., R_T^{*(i)}\right\}$ and used for the construction of one-step-ahead forecast of the volatility as follows

$$\widehat{s}_{T+1}^{*2(i)} = \widehat{\alpha}_0^{*(i)} + \widehat{\alpha}_1^{*(i)} R_T^2 + \widehat{\beta}^{*(i)}\widehat{s}_T^{*(i)} \tag{4.9}$$

where $\widehat{s}_T^{*2(i)} = \frac{\widehat{\alpha}_0^{*(i)}}{1-\widehat{\alpha}_1^{*(i)}-\widehat{\beta}^{*(i)}} + \widehat{\alpha}_1^{*(i)} \sum_{j=0}^{T-2} \widehat{\beta}^{*j(i)} \left(R_{T-j-1}^2 - \frac{\widehat{\alpha}_0^{*(i)}}{1-\widehat{\alpha}_1^{*(i)}-\widehat{\beta}^{*(i)}}\right)$, so that the forecast $\widehat{s}_{T+1}^{*2(i)}$ is based on the original series of returns $\{R_1, ..., R_T\}$ and on the bootstrap parameters.

Christoffersen and GonÇalves (2005) propose to compute prediction intervals for the $VaR_{T+1}$ and $ES_{T+1}$ by obtaining bootstrap replicates of the $VaR$ and $ES$ by the following expressions

$$\widehat{VaR}_{T+1}^{*(i)} = \widehat{s}_{T+1}^{*(i)}\widehat{q}^{*(i)} \tag{4.10}$$

and

$$\widehat{ES}_{T+1}^{*(i)} = \widehat{s}_{T+1}^{*(i)}\widehat{E}^{*(i)}, \tag{4.11}$$

where $\widehat{s}_{T+1}^{*(i)}$ is given by (4.9). They consider several alternative estimators of $q$ and $\underset{t-1}{E}\left[\epsilon_t \mid \epsilon_t \leq q\right]$. First, they assume that $\epsilon_t$ has a Normal distribution; see also Hartz and Paolella (2006) who propose a modification to avoid biases in the estimates of the $VaR$. In this case, $\widehat{q}^*$ and $\widehat{E}^*$ are given by

$$\widehat{q}^* = \Phi_{0.01}^{-1} \tag{4.12}$$

$$\widehat{E}^* = -\frac{\phi\left(\Phi_{0.01}^{-1}\right)}{0.01} \tag{4.13}$$

where $\Phi_{0.01}^{-1}$ is the quantile of the standardized Normal distribution and $\phi$ is its density function. Then, the values in (4.12) and (4.13) are substituted in (4.10) and (4.11) in order to obtain the corresponding bootstrap estimates $\widehat{VaR}_{T+1}^{*(i)}$ and $\widehat{ES}_{T+1}^{*(i)}$. Consequently, a set of B bootstrap estimates are obtained for both measures $\left\{\widehat{VaR}_{T+1}^{*(1)}, ..., \widehat{VaR}_{T+1}^{*(B)}\right\}$ and $\left\{\widehat{ES}_{T+1}^{*(1)}, ..., \widehat{ES}_{T+1}^{*(B)}\right\}$. The empirical distributions of $\widehat{VaR}_{T+1}^{*(i)}$ and $\widehat{ES}_{T+1}^{*(i)}$ are denoted by $Q_V^*(r) = \frac{\#\left\{\widehat{VaR}_{T+1}^{*(i)} \leq r\right\}}{B}$ and $Q_E^*(r) = \frac{\#\left\{\widehat{ES}_{T+1}^{*(i)} \leq r\right\}}{B}$ respectively, where $\#\{\cdot\}$ is the cardinality of $\{\cdot\}$. Then, the bootstrap prediction intervals for the $VaR_{T+1}$ and $ES_{T+1}$ are given by

$$\left[q_{\frac{\gamma}{2}}\left(Q_V^*(r)\right), q_{1-\frac{\gamma}{2}}\left(Q_V^*(r)\right)\right] \tag{4.14}$$

$$\left[q_{\frac{\gamma}{2}}\left(Q_E^*(r)\right), q_{1-\frac{\gamma}{2}}\left(Q_E^*(r)\right)\right]. \tag{4.15}$$

respectively, where $q_\gamma(\cdot)$ is the $\gamma th$ empirical quantile of the corresponding empirical distribution.

Alternatively, assuming that $\epsilon_t$ has a standardized Student-$\nu$ distribution, where $\nu$ are the degrees of freedom, the expression for the quantile and the expectation are given by

$$\widehat{q}^* = \sqrt{\frac{\nu-2}{\nu}} t_{0.01}^{-1} \tag{4.16}$$

$$\widehat{E}^* = \frac{\nu-2}{\alpha(1-\nu)} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi(\nu-2)\Gamma\left(\frac{\nu}{2}\right)}} \left(1+\frac{\widehat{q}^{*2}}{\nu-2}\right)^{\frac{1-\nu}{2}} \tag{4.17}$$

where $t_{0.01}^{-1}$ is the 0.01 quantile of the Student-$\nu$ distribution and $\Gamma(\cdot)$ is the gamma function. Then, as described above, these values are substituted in (4.10) and (4.11) respectively to compute $\widehat{VaR}_{T+1}^{*(i)}$ and $\widehat{ES}_{T+1}^{*(i)}$ and their respective prediction intervals. It is important to point out that the prediction intervals of the one-step ahead $VaR$ and $ES$ computed when assuming a particular error distribution incorporates the uncertainty associated with the estimation of the conditional variance but not that due to the estimation of the quantile and the expectation. Christoffersen and GonÇalves (2005) also propose to estimate $q$ and $\underset{t-1}{E}[\epsilon_t \mid \epsilon_t \leq q]$ in the $EVT$ framework by assuming that the tail of the conditional distribution of $\widehat{\epsilon}_t^*$ is well approximated by the distribution proposed by Gnedenko (1943). Then, the following Hill estimator is implemented to estimate the quantile of the $ith$ bootstrap replicate

$$\widehat{q}^{*(i)} = -\widehat{\epsilon}^{*(i)}_{(k+1)} \left(\frac{0.01}{k/T}\right)^{-\widehat{\xi}^{(k)(i)}} \tag{4.18}$$

where $k << T$ is the number of observations in the tail and $\widehat{\epsilon}^{*(i)}_{(k+1)}$ is the $(k+1)\, th$ order statistic of the bootstrap residuals which are given by

$$\widehat{\epsilon}^{*(i)}_t = \frac{R^{*(i)}_t}{\widehat{\sigma}^{*(i)}_t} \tag{4.19}$$

where $\widehat{\sigma}^{*2(i)}_t = \widehat{\alpha}^{*(i)}_0 + \widehat{\alpha}^{*(i)}_1 R^{*2(i)}_{t-1} + \widehat{\beta}^{*(i)}_1 \widehat{\sigma}^{*2(i)}_{t-1}$. It is important to note the difference between $\widehat{s}^{*2(i)}_{T+1}$ as defined in (4.12) and $\widehat{\sigma}^{*2(i)}_t$. The latter is calculated using the bootstrap replicates $\left\{R^{*(i)}_1, ..., R^{*(i)}_T\right\}$ while the former uses the original observations $\{R_1, ..., R_T\}$. Finally, $\widehat{\xi}^{(k)(i)} = \dfrac{1}{k}\sum_{j=1}^{k} \log\left(\widehat{\epsilon}^{*}_{(j)}\right) - \log\left(\widehat{\epsilon}^{*}_{(k+1)}\right)$.

On the other hand, the expectation in the tail is given by

$$\widehat{E}^{*(i)} = \frac{\widehat{q}^{*(i)}}{1-\widehat{\xi}^{(k)(i)}}. \tag{4.20}$$

where $\widehat{q}^{*(i)}$ is given in $(4.18)$. Implementing the Hill estimator requires to choose the cut-off point $k$ which defines the sub-sample of extremes from which the tail index parameter, $\xi$, is estimated. However, there is not a formal method proposed in the literature in order to choose the optimal value of $k$, but . Consequently, Christoffersen and GonÇalves (2005) compute the bias and root mean squared error of the $\widehat{VaR}^{*}_{T+1}$ and $\widehat{ES}^{*}_{T+1}$ and choose $k$ as the value for which they are stable. Alternatively, in this Chapter, we choose the value of $k$ such that $\widehat{\xi}^{(k)}$ is stable.

As before, once $q$ and $\underset{t-1}{E}\left[\epsilon_t \mid \epsilon_t \leq q\right]$ are estimated, they are substituted in (4.10) and (4.11) to obtain bootstrap replicates of $\widehat{VaR}^{*}_{T+1}$ and $\widehat{ES}^{*}_{T+1}$. The corresponding intervals will be denoted as $CG-H$.

Christoffersen and GonÇalves (2005) also propose two estimators of $q$ and $\underset{t-1}{E}\left[\epsilon_t \mid \epsilon_t \leq q\right]$ that do not rely on any particular error distribution. In particular, the $FHS$ consists on estimating $q$ by the $\omega th$-order statistic of the standardized bootstrap residuals

$$\widehat{q}^{*(i)} = \widehat{\epsilon}^{*(i)}_{(\omega)} \tag{4.21}$$

where $\omega = [T \times 0.01] = \max \{m \mid m \leq T \times 0.01, m \in \mathbb{N}\}$. The corresponding expectation is estimated by

$$\widehat{E}^{*(i)} = \frac{\sum_{j=1}^{\omega} \epsilon_{(j)}^{*(i)}}{\omega}. \tag{4.22}$$

The corresponding intervals will be denoted as $CG - FHS$.

Finally, they propose using the Gram-Charlier and Cornish-Fisher ($GCCF$) expansions to approximate the conditional density of the standardized bootstrap residuals. The Cornish-Fisher expansion is used for the estimation of $q$ as follows

$$\widehat{q}^{*(i)} = \Phi_{0.01}^{-1} + \frac{\widehat{\gamma}_1}{6}\left[\left(\Phi_{0.01}^{-1}\right)^2 - 1\right] + \frac{\widehat{\gamma}_2}{24}\left[\left(\Phi_{0.01}^{-1}\right)^3 - 3\Phi_{0.01}^{-1}\right] - \frac{\widehat{\gamma}_1^2}{36}\left[2\left(\Phi_{0.01}^{-1}\right)^3 - 5\Phi_{0.01}^{-1}\right] \tag{4.23}$$

where $\widehat{\gamma}_1 = \frac{1}{T}\sum_{t=1}^{T} \widehat{\epsilon}_t^{*3(i)}$ and $\widehat{\gamma}_2 = \frac{1}{T}\sum_{t=1}^{T} \widehat{\epsilon}_t^{*4(i)} - 3$. Giamouridis (2006) provides the following correction for the expression of the expectation needed to compute the $ES$

$$\widehat{E}^{*} = \frac{\phi\left(\widehat{q}^{*(i)}\right)}{0.01}\left(1 + \frac{\widehat{\gamma}_1}{6}\left(\widehat{q}^{*(i)}\right)^3 + \frac{\widehat{\gamma}_2}{24}\left[\left(\widehat{q}^{*(i)}\right)^4 - 2\left(\widehat{q}^{*(i)}\right)^2 - 1\right]\right) \tag{4.24}$$

where $\widehat{q}^{*(i)}$ is given as in (4.23). The corresponding intervals are denoted as $CG - GCCF$.

They show that the bootstrap intervals constructed for the $FHS$ estimator of $VaR_{T+1}$ have coverages close to the nominal. However, the coverages of the bootstrap $FHS$ estimator of $ES_{T+1}$ are well under the nominal. On the other hand, the bootstrap intervals of the Hill estimator of $ES_{T+1}$ are closer but still well under the nominal coverage. Therefore, the prediction intervals cannot be trusted for the $ES$ risk measures.

Recently, Ho and Lee (2005) show that the traditional bootstrap procedures are not adequate when constructing prediction intervals for quantiles. Consequently, they propose the iterated smoothed bootstrap which corrects the errors in estimating quantiles by calibrating the nominal coverage level iteratively while smoothing the bootstrap amounts to drawing bootstrap samples from a kernel-smoothed empirical distribution instead of sampling with replacement from the row data. Ho and Lee (2005) assume that the original data is iid. Therefore, their procedure can be implemented, in our case, to the standardized returns $\widehat{\epsilon}_t$, obtaining then, a confidence interval for $q$ which allow us to construct a confidence interval for the one-step ahead $VaR$ by multiplying it by $\widehat{\sigma}_{T+1}$. Next, we describe the iterated smoothed bootstrap.

If $F$ is the distribution of standardized returns, $F^{-1}(\alpha)$ is the $\alpha th$ quantile of $\widehat{\epsilon}_t$. Let $F_T$ be the empirical distribution function of $\widehat{\epsilon}_t$ and $\widehat{F}_{T,\eta}(t) = T^{-1}\sum_{i=1}^{T} K((t-\widehat{\epsilon}_i)/\eta)$ its kernel-smoothed version defined for a kernel function $K$ and a bandwidth $\eta > 0$. Ho and Lee (2005) propose the triangular kernel given by $k(x) = 1-|x|$, for $|x| \leq 1$. With respect to the selection of an optimal bandwidth, they propose a bootstrap procedure. The idea is to fix a grid of proposed bandwidths, then generate smoothed and smoothed iterated intervals for each bandwidth and selecting the bandwidth which gives the best coverage. This procedure is very time consuming because of the use of many replicates bootstrap in each case. This is even worse for the smoothed bootstrap because for each bootstrap replicate $B$, a bootstrap replicate $C$ is generated. Therefore, we need a lot of time for selecting the bandwidth and then also for constructing the interval.

The smoothed bootstrap percentile method is the following: let $\widehat{\epsilon}_t^{\dagger} = \left\{\widehat{\epsilon}_1^{\dagger}, ...\widehat{\epsilon}_T^{\dagger}\right\}$ be a random sample from $\widehat{F}_{T,\eta}$ which in practice is generated by $\widehat{\epsilon}_i^{\dagger} = Y_i^* + \eta W_i^*$, where $Y_i^*$ and $W_i^*$ are independent random draws from $F_T$ and $K$ respectively. Let $F_{T,\eta}^*$ be the empirical distribution of $\widehat{\epsilon}_t^{\dagger}$ and define

$$G_T(t) = P\left[T^{1/2}\left(F_T^{-1}(\alpha) - F^{-1}(\alpha)\right) \leq t\right], t \in \mathbb{R}. \tag{4.25}$$

The smoothed version of (28) is given by

$$\widehat{G}_{T,\eta}(t) = P\left[T^{1/2}\left(F_{T,\eta}^{*-1}(\alpha) - \widehat{F}_{T,\eta}^{-1}(\alpha)\right) \leq t \mid \widehat{\epsilon}_t\right], t \in \mathbb{R}. \tag{4.26}$$

A noniterated smoothed bootstrap confidence interval for $F^{-1}(\alpha)$ is given by

$$I_{1,\gamma} = \left[-\infty, F_T^{-1}(\alpha) - T^{-1/2}\widehat{G}_{T,\eta}^{-1}(\gamma)\right]. \tag{4.27}$$

Consequently, the noniterated smoothed version for the $VaR_{T+1}$ would be

$$\widehat{\sigma}_{T+1} I_{1,\gamma}. \tag{4.28}$$

The next step is to iterate the smoothed bootstrap. Let $\widehat{\epsilon}_t^{\dagger *} = \left\{\widehat{\epsilon}_1^{\dagger *}, ..., \widehat{\epsilon}_T^{\dagger *}\right\}$ be a generic outer-level random sample from $\widehat{F}_{T,\beta}$, for a bandwidth $\beta > 0$ and $F_{T,\beta}^*$ be its empirical distribution function. Define the smoothed empirical distribution of $\widehat{\epsilon}_t^{\dagger *}$ as $\widehat{H}_{T,\eta}(t) = T^{-1} \sum_{i=1}^{T} K\left(\left(t - \widehat{\epsilon}_i^{\dagger *}\right)/\eta\right)$. Denote by $\widehat{\epsilon}_t^{**} = \{\widehat{\epsilon}_1^{**}, ..., \widehat{\epsilon}_T^{**}\}$ a generic inner-level sample drawn from $\widehat{H}_{T,\eta}$ and by $H_{T,\eta}^{*}$ its empirical distribution. Define

$$\widehat{G}_{T,\eta}^{*}(t) = P\left[T^{1/2}\left(H_{T,\eta}^{*-1}(\alpha) - \widehat{H}_{T,\eta}^{-1}(\alpha)\right) \leq t \mid \widehat{\epsilon}_t, \widehat{\epsilon}_t^{\dagger *}\right], t \in \mathbb{R}. \tag{4.29}$$

Then, using $\widehat{J}_{n,\beta,\eta}$, the conditional distribution of $\widehat{G}_{T,\eta}^{*}\left(T^{1/2}\left(F_{T,\beta}^{*-1}(\alpha) - \widehat{F}_{T,\beta}^{-1}(\alpha)\right)\right)$ given $\widehat{\epsilon}_t$, the iterated smoothed bootstrap confidence interval for $F^{-1}(\alpha)$ is obtained by the following expression

$$I_{2,\gamma} = \left[-\infty, F_T^{-1}(\alpha) - T^{-1/2}\widehat{G}_{T,\eta}^{-1}(\widehat{J}_{n,\beta,\eta}(\gamma))\right]. \tag{4.30}$$

Consequently, the iterated smoothed version for the $VaR_{T+1}$ would be

$$\widehat{\sigma}_{T+1} I_{2,\gamma}. \tag{4.31}$$

### 4.2.2 A new bootstrap procedure

The bootstrap procedures proposed by Christoffersen and GonÇalves (2005) are based on estimating $q$ and $\underset{t-1}{E}[\epsilon_t \mid \epsilon_t \leq q]$ from the bootstrap standardized returns given by $\widehat{\epsilon}_t^{*} = \frac{R_t^{*}}{\widehat{\sigma}_t^{*}}$. However, note that when computing the bootstrap replicates as in (4.8), the bootstrap conditional variances depend on the QML estimates and the bootstrap returns; see (4.7). On the other hand, the estimated conditional standard deviations used to standardized the returns, $\widehat{\sigma}_t^{*(i)}$, depend on bootstrap estimates of the parameters and bootstrap returns; see (4.19). Therefore, the standardized bootstrap returns used by Christoffersen and GonÇalves (2005) are given by $\epsilon_t^{*}\frac{\sigma_t^{*}}{\widehat{\sigma}_t^{*}}$, which depends on the ratio $\frac{\sigma_t^{*}}{\widehat{\sigma}_t^{*}}$. In this Chapter, we propose to simplify the procedure from a methodological point of view by estimating the quantiles and the expectations required to compute the $VaR$ and $ES$ using the residuals bootstrapped from the original residuals, i.e, $\epsilon_t^{*}$. In this way, the statistical properties of the procedure are much easier to prove and it is simplified as we do not need to compute the residuals $\widehat{\epsilon}_t^{*}$. Furthermore, in small sample sizes the quantiles and expectations, estimated as proposed by Christoffersen and GonÇalves (2005), are computed using $T$ observations which is well known to have not adequate properties for small $T$. However, we can obtain as many bootstrap replicates of $\epsilon_t^*$ as desired, so that the quantile and the expectation in the tail are computed using a large number of observations in the tail.

Next, we describe the proposed bootstrap algorithm to obtain prediction intervals for $\widehat{VaR}^*_{T+1}$ and $\widehat{ES}^*_{T+1}$, which are based on a second bootstrap step without making any particular assumption on the error distribution.

For each bootstrap replicate of the series of returns, $\widehat{s}^{*2}_{T+1}$ is obtained as in equation (4.9). Then, we obtain $\epsilon^{*(i,n)}_{T+1}$ random draws from the empirical distribution of the standardized residuals $\widehat{\epsilon}_t$, for $n = 1, ..., N$. Therefore, for each bootstrap replicate of the original series of returns, $i = 1, ..., B$, we obtain a set of $N$ disturbances $\left\{\epsilon^{*(i,1)}_{T+1}, ..., \epsilon^{*(i,N)}_{T+1}\right\}$. The quantile, $q$, and the expectation in the tail can then be calculated by any of the following three alternatives. First, using FHS, they can be obtained as the 1% quantile of the empirical distribution of $\epsilon^{*(i,n)}_{T+1}$ given by $F^*_B(\epsilon) = \dfrac{\#\left\{\epsilon^{*(i,n)}_{T+1} \leq \epsilon\right\}}{N}$. Therefore,

$$\widehat{q}^{*(i)} = \widehat{\epsilon}^{*(i,n)}_{(\omega)} \tag{4.32}$$

$$\widehat{E}^{*(i)} = \frac{\sum_{j=1}^{\omega} \epsilon^{*(i,n)}_{(j)}}{\omega}$$

where $\omega$ is as defined before. After obtaining $B$ measures of risk given by $\left\{\widehat{VaR}^{*(1)}_{T+1}, ..., \widehat{VaR}^{*(B)}_{T+1}\right\}$ and $\left\{\widehat{ES}^{*(1)}_{T+1}, ..., \widehat{ES}^{*(B)}_{T+1}\right\}$. Finally, the $100(1-\gamma)\%$ prediction interval for the $VaR$ and $ES$ can be obtained by the percentile method of Efron (1981, 1982) given by[1]

$$\left[q_{\frac{\gamma}{2}}\left(Q^*_{VB}(r)\right), q_{1-\frac{\gamma}{2}}\left(Q^*_{VB}(r)\right)\right] \tag{4.33}$$

$$\left[q_{\frac{\gamma}{2}}\left(Q^*_{EB}(r)\right), q_{1-\frac{\gamma}{2}}\left(Q^*_{EB}(r)\right)\right]. \tag{4.34}$$

[1] In order to improve the coverage of the confidence intervals Efron (1981, 1982) propose the bias corrected method that applied to the $VaR$ case is given by $\left[Q^{*-1}_{VB}\left(\Phi\left[2z_0 + z_{\frac{\gamma}{2}}\right]\right), Q^*_{VB}\left(\Phi\left[2z_0 + z_{1-\frac{\gamma}{2}}\right]\right)\right]$, where $z_\gamma$ is the $\gamma th$ quantile of a Normal distribution and $z_0 = \Phi^{-1}\left[Q^*_{VB}\left(\widehat{VaR}_{T+1}\right)\right]$. Similar formulation is given by the $ES$ case. Using this procedure we obtain the same conclusions than with the percentile method. Alternatively, Efron (1987) proposes a generalization of the bias corrected method named, the accelerated bias corrected method. Berkowitz and Kilian (2000) mention that the implementation of the latter method to time series has not been investigated and it is not straightforward. Therefore, we do not consider this alternative.

where $Q^*_{VB}(r)$ and $Q^*_{EB}(r)$ are the bootstrap empirical distribution function of $\widehat{VaR}^{*(i)}_{T+1}$ and $\widehat{ES}^{*(i)}_{T+1}$ respectively. The corresponding intervals are denoted as $NR-FHS$.

Alternatively, the quantile and expectation can also be estimated by using the Hill estimators in (4.18) and (4.20) implemented to the bootstrap standardized residuals $\epsilon^*_t$. The corresponding intervals will be denoted as $NR-H$. Finally, the $GCCF$ expansion in (4.23) and (4.24) could be implemented and the intervals denoted as $NR-GCCF$.

## 4.3 Monte Carlo experiments

In this section, we carry out Monte Carlo experiments to analyze the finite sample performance of the proposed bootstrap procedure for constructing prediction intervals for the $VaR$ and $ES$. We compare our procedure with those proposed by Christoffersen and GonÇalves (2005) and Ho and Lee (2005).

We generate 1000 replicates from a $GARCH(1,1)$ model with parameters $\alpha_0 = 0.002$, $\alpha_1 = 0.05$ and $\beta = 0.9$. In order to take into account the potential presence of leverage effect, we also simulate series using a $EGARCH(1,1)$ model with $\alpha_0 = -0.17$, $\alpha_1 = 0.15$, $\gamma = -0.13$ and $\beta = 0.9$[2]. We consider four sample sizes, $T = 250, 500, 1000$ and 3000 and three alternative distributions of the standardized observations, $\epsilon_t$, namely Normal, Student-8 and a Skewed-Student distribution, with 10 degrees of freedom and a coefficient of asymmetry equal to $-0.11$; see Hansen (1994) for a complete description of the Skewed distribution. Although $T = 250$ may seem a rather small value for real life applications, it is important to note that the Basel Commission requires to compute the $VaR$ with at least one year of data which corresponds approximately to 250 daily observations. In each case, we estimate the parameters of the True Data Generating Process by $ML$ by treating the degrees of freedom and asymmetry parameters as unknown parameters. We compute 90% prediction intervals for the 1% $VaR_{T+1}$ and $ES_{T+1}$ based on $B = 1000$ bootstrap replicates by implementing the

[2]The Exponential GARCH ($EGARCH$) model of Nelson (1991) is given by
$\ln\left(\sigma^2_t\right) = \alpha_0 + \alpha_1\left[|\epsilon_{t-1}| - E\left(|\epsilon_{t-1}|\right)\right] + \beta_1 \ln\sigma^2_{t-1} + \gamma\epsilon_{t-1}$

procedures proposes by Christoffersen and GonÇalves (2005) by assuming i) the true distribution ($C-G-D$), ii) the $FHS$ ($C-G-FHS$), iii) $EVT$ ($C-G-H$) and iv) Gram-Charlier and Cornish-Fisher expansions ($C-G-GCCF$). We also construct the prediction intervals by the bootstrap procedure proposed in this Chapter by i) $FHS$ ($NR-FHS$), ii) $EVT$ ($NR-H$) and iii) by the Gram-Charlier and Cornish-Fisher approximations ($NR-GCCF$). Finally, the prediction intervals are computed by the procedure proposed by Ho and Lee (2005) ($HL$) using $B = 1000$ bootstrap replicates, and for the smoothed iterated, $B = 1000$ replicates for the outer level and $C = 500$ replicates for the inner level.

Table 4.1, that reports the coverages of the 90% intervals[3] for the $VaR_{T+1}$ and $ES_{T+1}$ when the series are generated by the $GARCH(1,1)$ model, shows that regardless of the particular distribution, the coverages of the $VaR$ computed by assuming the true error distribution are well under the nominal. The same can be said when using the Hill estimator regardless of whether the bootstrap replicates are obtained as proposed by Christoffersen and GonÇalves (2005) or as proposed in this Chapter. It seems that assuming a particular error distribution does not account for the uncertainty associated with the estimation of the $VaR$ even when the assumed distribution is the true one. The coverages are closer to the nominal when using either $FHS$ or the $GCCF$ expansions and slightly better with the former. Furthermore note that the coverages are clearly closer to the nominal 90% when using the bootstrap procedure proposed in this Chapter instead of that proposed by Christoffersen and GonÇalves (2005). Therefore, with respect to the one-step ahead $VaR$, the bootstrap procedure proposed in this Chapter implemented to $FHS$ gives coverages closer to the nominal for all assumed distributions and sample sizes. Christoffersen and GonÇalves (2005) also conclude that $FHS$ gives the best results when implemented using their bootstrap procedure. Furthermore, note that the coverages obtained using the procedure proposed in this Chapter are comparable to those obtained when the much more complicated procedure of Ho and Lee (2005) is implemented.

[3]The conclusions based on results for 95% intervals are similar and not reported to save space.

Additionally, Table 4.1 also reports the average coverages for the $ES$. In this case, we can observe that the intervals constructed assuming the true distribution are once more well under nominal coverage. Furthermore, when comparing the coverages obtained when implementing $FHS$, Hill estimator or the $GCCF$ expansions, it seems that the latter is, in general more appropriate regardless of whether one uses the bootstrap procedure proposed by Christoffersen and GonÇalves (2005) or the one proposed in this Chapter. Although this result seems to contradict the conclusions of Christoffersen and GonÇalves (2005), notice that we are implementing the correction in the estimation of the expectation of the tail proposed by Giamouridis (2006). As a consequence, the performance of the prediction intervals of one-step ahead ES constructed using the GCCF expansion is clearly improved. When comparing the results obtained when using our bootstrap procedure with those obtained using $\widehat{\epsilon}_t^*$, we can observe that the coverages are in general closer to the nominal when implementing the second bootstrap step proposed in this Chapter. In any case, our bootstrap intervals suffer from a slight overcoverage which is preferable from a conservative risk management strategy than to have undercoverage.

Table 4.2 reports the coverages of the 90% intervals for the $VaR_{T+1}$ and $ES_{T+1}$ when the series are generated by the $EGARCH$ model. The first clear difference between the results obtained for the $GARCH$ and the $EGARCH$ models is that in the former, the coverages are smaller than the nominal while in the latter, there is an overcoverage. In any case, although the differences between alternative procedures are smaller than those observed in Table 4.1 for the $GARCH$ model, we still observe that the coverages are closer to the nominal when using the two-step bootstrap proposed in this Chapter. In the context of the $EGARCH$ model, $FHS$ seems to work better for both $VaR$ and $ES$ prediction intervals.

On the other hand, last column of Table 4.1 reports the coverage of the procedure of Ho and Lee (2005) using the $GARCH(1,1)$ model. We can observe, that the coverages constructed with the smoothed bootstrap are better than those obtained with the percentile method, except when the data is generated using the Skewed-Student distribution and the sample size is 250. Similar results can be concluded from the last column of Table 4.2 when comparing the coverages obtained with the $EGARCH$ model. In this case, for all de distributions and sample sizes the coverages are improved.

## 4.4 Empirical Application

In this section, we implement the methods described above to obtain prediction densities for $VaR$ and $ES$ of three series of daily returns, the $S\&P500$ index, the $IBEX35$ and the $Euro/Dollar$ observed from 01/09/2003 to 19/01/2010 which have been downloaded from the EcoWin database. Figure 4.1, that plots the series of returns, shows the effect of the crisis on an increased volatility at the end of the sample period. Figure 4.1 also plots for each of the three series, the correlogram of absolute returns and the cross-correlations between returns, $y_t$, and future squared returns, $y_{t+h}^2$ together with their 95% confidence bands computed as suggested by Diebold (1988) to account for the presence of conditional heteroscedasticity. It is clear that the sample correlations of absolute returns are positive and highly persistent, being significantly different from zero even for very long lags. Therefore, returns could be conditionally heteroscedastic, possibly with long-memory. In the case of the $Euro/Dollar$ the sample correlations of absolute returns are also positive but the presence of conditional heteroscedasticity is less strong. However, the cross-correlations do not give any clear evidence of the presence of leverage effect.

Table 4.3 shows the 90% prediction intervals for the one-step ahead $VaR$ and the $ES$ constructed at the end of the sample, constructed using the procedures described in Section 4.2 using the $GARCH(1,1)$ and the $EGARCH(1,1)$ models with Normal errors for the $S\&P500$, the $IBEX35$ and the $Euro/Dollar$ returns. We observe that the intervals for the $VaR$ are clearly narrower than those of the $ES$ and the intervals using the $GARCH(1,1)$ model are also narrower than those using the $EGARCH(1,1)$. In the Monte Carlo experiments we observe the same behavior because the coverages using the $EGARCH(1,1)$ were over the nominal. We can also notice that the differences among interval widths is smaller in the case of the $Euro/Dollar$. This could be expected because of the lack of heteroscedaticity. In the case of the $S\&P500$, the $IBEX35$ the differences among procedures are remarkable.

The intervals for the $VaR$ calculated using the $GCCF$ approximations are clearly wider than the others when using the bootstrap procedure proposed in this Chapter and also with that proposed by Christoffersen and GonÇalves (2005). On the other hand, the intervals constructed with the Normal distribution are the narrowest. Moreover, regardless the model, when implementing $FHS$, Hill estimator or the $GCCF$ expansions the widths are larger for the $ES$. This is consistent with the results obtained in the Monte Carlo experiment where the best coverages were obtained when using these procedures.

We can also notice that the $VaR$ prediction interval widths obtained with the procedure proposed in this Chapter using $FHS$ are not always larger than the rest of the procedures. However, as we conclude in the last section, it provides the best coverages.

With respect to the intervals calculated using the procedure of Ho and Lee (2005) we can observe that the upper limits are greater than those obtained with the rest of the procedures when using the $GARCH(1,1)$ model for the $S\&P500$, the $IBEX35$. On the other hand, for the $IBEX35$ the upper limit is near to the others. However, when using the $EGARCH(1,1)$ model, the upper limit is, in general, lower than the others, except when the intervals are calculated using the $GCCF$ expansions. As long as the upper limit of the prediction intervals using the procedure of Ho and Lee (2005) are greater than the rest, the coverages will be closer to the nominal.

## 4.5 Conclusions

In this Chapter, we propose a new bootstrap procedure to construct prediction intervals for two of the most famous measures of risk, the $VaR$ and the $ES$. We propose a second bootstrap step which avoid an extra source of uncertainty by bootstrapping directly from the original standardized residuals when computing the quantile of the error distribution and the expectation in the tail.

Furthermore, we incorporate in our procedure the known fact that the financial series of returns have leverage effect. Thus, we calculate the volatility assuming the $EGARCH$ model of Nelson (1991).

Several Monte Carlo experiments are carried out to analyze the finite sample properties of the proposed procedure and compare them with those of several alternatives. We show that in most of the cases, our procedure produces coverages closer to the nominal. The proposed procedure has coverages very close to the nominal when computing the VaR using the corresponding quantile of the bootstrap empirical distribution. However, when computing the ES the results can be better when using the Hill or $GCCF$ approximations.

For an illustration, we implement the proposed procedure to real series of returns, constructing the point estimates of the measures or risk and also the prediction intervals.

The extension of the analysis when the evolution of the volatility is modeled by using Stochastic Volatility $(SV)$ instead of $GARCH$ models is left for further research.

It will also be interesting to know which is the behavior of the proposed procedure in the presence of misspecification.

Table 4.1. GARCH(1,1) 90% Prediction intervals coverage rates obtained by Monte Carlo simulations

| 90% Prediction Intervals for the one-step-ahead 1%VaR | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Distribution | T | C-G-D | C-G-FHS | C-G-H | C-G-GCCF | NR-FHS | NR-H | NR-GCCF | Smoothed |
| Normal | 250 | 76.80% | 81.70% | 73.00% | 80.10% | **88.30%** | 72.90% | 83.80% | **91.40%** |
| | 500 | 72.30% | 83.30% | 73.60% | 79.20% | **87.40%** | 73.30% | 83.40% | **91.60%** |
| | 1000 | 71.40% | 80.60% | 74.40% | 79.70% | **84.20%** | 75.90% | 83.50% | **90.90%** |
| | 3000 | 70.30% | 80.60% | 72.20% | 79.10% | **84.70%** | 72.10% | 83.10% | **90.60%** |
| Student-v | 250 | 77.10% | 82.80% | 74.50% | 76.40% | **88.70%** | 72.30% | 82.20% | **88.70%** |
| | 500 | 74.30% | 86.17% | 76.83% | 80.02% | **89.01%** | 76.60% | 84.16% | **89.10%** |
| | 1000 | 73.00% | 84.70% | 76.60% | 81.10% | **87.80%** | 77.50% | 87.20% | **91.50%** |
| | 3000 | 78.30% | 86.30% | 83.20% | 77.70% | **89.30%** | 84.20% | 84.70% | **89.40%** |
| Skewed Student-v | 250 | 76.40% | 84.20% | 73.50% | 77.40% | **89.50%** | 71.40% | 84.70% | **91.90%** |
| | 500 | 76.20% | 87.30% | 79.40% | 82.10% | **89.60%** | 78.90% | 87.70% | **89.90%** |
| | 1000 | 75.50% | 87.10% | 78.10% | 82.80% | **89.50%** | 79.90% | 87.80% | **90.60%** |
| | 3000 | 76.60% | 86.40% | 81.70% | 81.20% | **89.70%** | 83.30% | 87.70% | **90.20%** |

| 90% Prediction Intervals for the one-step-ahead 1%ES | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Distribution | T | C-G-D | C-G-FHS | C-G-H | C-G-GCCF | NR-FHS | NR-H | NR-GCCF |
| Normal | 250 | 76.80% | 70.90% | 80.90% | **87.80%** | 74.40% | 84.90% | 86.20% |
| | 500 | 72.30% | 73.90% | 82.40% | 88.10% | 76.00% | 86.50% | **88.70%** |
| | 1000 | 71.40% | 76.80% | 81.40% | 84.20% | 79.30% | 88.80% | **90.00%** |
| | 3000 | 70.30% | 78.60% | 79.10% | 80.60% | 80.60% | 86.90% | **88.00%** |
| Student-v | 250 | 77.10% | 63.30% | 77.80% | **93.40%** | 67.30% | 80.10% | 97.90% |
| | 500 | 74.30% | 70.09% | 82.27% | **87.12%** | 71.28% | 72.60% | 96.45% |
| | 1000 | 73.00% | 77.20% | 84.50% | 85.50% | 78.30% | **87.30%** | 96.60% |
| | 3000 | 78.30% | 84.60% | 86.60% | 88.40% | 85.70% | **91.40%** | 94.60% |
| Skewed Student-v | 250 | 76.40% | 60.50% | 78.00% | **89.90%** | 65.60% | 80.90% | 94.50% |
| | 500 | 76.20% | 75.20% | 84.50% | 86.00% | 75.70% | 85.50% | **92.60%** |
| | 1000 | 75.50% | 81.30% | 87.60% | 86.20% | 82.90% | **90.40%** | 95.10% |
| | 3000 | 76.60% | 81.70% | 85.30% | 77.90% | 83.90% | **89.10%** | 91.50% |

Table 4.2. EGARCH(1,1) 90% Prediction intervals coverage rates obtained by Monte Carlo simulations

| 90% Prediction Intervals for the one-step-ahead 1%VaR | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Distribution | T | C-G-D | C-G-FHS | C-G-H | C-G-GCCF | NR-FHS | NR-H | NR-GCCF | Smoothed |
| Normal | 250 | 93.10% | 93.30% | 94.40% | 93.40% | **92.60%** | 93.60% | 94.70% | **88.90%** |
| | 500 | 93.90% | 93.90% | 93.80% | 94.70% | **93.80%** | 93.70% | 94.30% | **87.90%** |
| | 1000 | 94.30% | 94.50% | 94.50% | 95.00% | **94.20%** | 94.70% | 94.50% | **87.80%** |
| | 3000 | **92.50%** | 93.10% | 93.20% | 93.10% | 93.00% | 93.20% | **92.50%** | **89.60%** |
| Student-v | 250 | 92.70% | 94.00% | 93.70% | 94.90% | **92.50%** | 92.70% | 93.70% | **90.50%** |
| | 500 | 93.40% | 94.80% | 93.40% | 94.50% | 93.80% | **93.30%** | 93.90% | **91.20%** |
| | 1000 | 90.70% | 91.10% | 90.90% | 91.00% | 90.70% | 90.80% | **90.60%** | **88.80%** |
| | 3000 | 91.60% | 91.70% | 91.70% | **91.50%** | 91.60% | **91.50%** | **91.50%** | **89.70%** |
| Skewed Student-v | 250 | 92.60% | 94.20% | 92.60% | 94.10% | 92.60% | **92.50%** | 93.30% | **90.70%** |
| | 500 | 93.40% | 93.50% | 94.00% | 94.10% | **92.70%** | 93.20% | 93.50% | **90.40%** |
| | 1000 | 92.10% | 92.40% | 91.40% | 92.60% | 91.90% | **91.70%** | 91.80% | **89.20%** |
| | 3000 | 88.60% | 88.60% | 88.20% | 88.70% | **88.80%** | 88.30% | **88.80%** | **89.80%** |

| 90% Prediction Intervals for the one-step-ahead 1%ES | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Distribution | T | C-G-D | C-G-FHS | C-G-H | C-G-GCCF | NR-FHS | NR-H | NR-GCCF |
| Normal | 250 | 93.10% | 94.00% | 94.20% | 91.60% | 93.40% | 95.50% | **91.60%** |
| | 500 | 93.90% | 94.10% | 94.40% | 94.00% | 94.20% | 94.90% | **93.50%** |
| | 1000 | 94.30% | **93.70%** | 94.00% | 94.30% | 94.00% | 94.10% | 94.10% |
| | 3000 | 92.50% | 92.90% | 92.50% | 93.00% | 92.50% | **92.40%** | **92.40%** |
| Student-v | 250 | 92.70% | 93.00% | 93.50% | 95.70% | **91.70%** | 93.10% | 94.90% |
| | 500 | 93.40% | 92.70% | 94.20% | 94.60% | **92.80%** | 93.40% | 94.20% |
| | 1000 | 93.50% | 93.00% | 93.80% | 92.90% | **92.80%** | 93.00% | 93.70% |
| | 3000 | 91.60% | 92.00% | **91.50%** | **91.50%** | 91.80% | 91.80% | **91.50%** |
| Skewed Student-v | 250 | 92.60% | 91.80% | 93.10% | 95.40% | **90.50%** | 93.20% | 94.10% |
| | 500 | 93.40% | 94.00% | 93.50% | 94.00% | **93.30%** | 94.00% | 94.40% |
| | 1000 | 92.10% | 92.10% | 92.50% | 92.70% | **92.00%** | 92.50% | 92.50% |
| | 3000 | 88.60% | 88.20% | 88.90% | 88.70% | 88.30% | **89.10%** | **89.10%** |

Figure 4.1. S&P500, IBEX35 and Euro/Dollar returns observed from 1th September 2003 up to 19th January 2010, correlogram of absolute returns and cross-correlogram of returns and squared returns together with their corresponding 95% confidence bands.

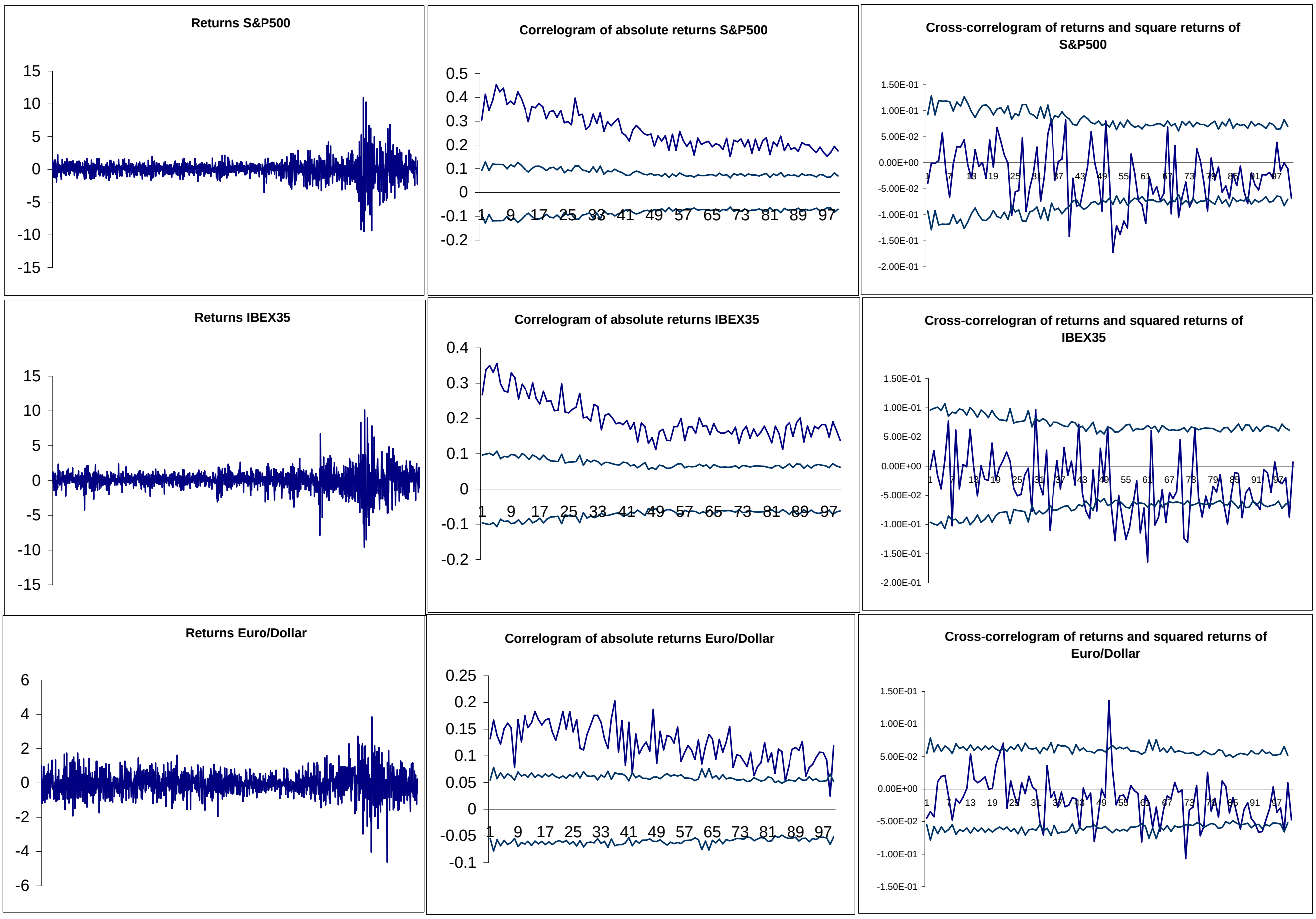


Table 4.3. 90% prediction intervals for 1% VaR and 1%ES for S&P500, IBEX35 and Euro/Dollar returns

| | | S&P500 | | | Euro/Dollar | | | IBEX35 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Model | | Lower Limit | Upper Limit | Width | Lower Limit | Upper Limit | Width | Lower Limit | Upper Limit | Width |
| GARCH-N | C-G-N | -0.0196 | -0.0178 | 0.0018 | -0.0146 | -0.0130 | 0.0016 | -0.0224 | -0.0205 | 0.0020 |
| | C-G-FHS | -0.0220 | -0.0194 | 0.0026 | -0.0164 | -0.0134 | 0.0030 | -0.0278 | -0.0222 | 0.0056 |
| | C-G-H | -0.0222 | -0.0192 | 0.0030 | -0.0155 | -0.0132 | 0.0023 | -0.0267 | -0.0224 | 0.0043 |
| | C-G-GCCF | -0.0278 | -0.0193 | 0.0085 | -0.0167 | -0.0137 | 0.0030 | -0.0322 | -0.0233 | 0.0089 |
| | NR-FHS | -0.0223 | -0.0191 | 0.0032 | -0.0166 | -0.0131 | 0.0035 | -0.0283 | -0.0217 | 0.0066 |
| | NR-H | -0.0223 | -0.0188 | 0.0035 | -0.0155 | -0.0129 | 0.0026 | -0.0270 | -0.0221 | 0.0049 |
| | NR-GCCF | -0.0286 | -0.0190 | 0.0097 | -0.0170 | -0.0135 | 0.0035 | -0.0336 | -0.0229 | 0.0107 |
| | H-L | | -0.0183 | 0.0183 | | -0.0135 | 0.0135 | | -0.0204 | 0.0204 |
| EGARCH-N | C-G-N | -0.0455 | -0.0118 | 0.0337 | -0.0240 | -0.0085 | 0.0155 | -0.0453 | -0.0146 | 0.0307 |
| | C-G-FHS | -0.0519 | -0.0132 | 0.0388 | -0.0255 | -0.0089 | 0.0165 | -0.0547 | -0.0171 | 0.0376 |
| | C-G-H | -0.0507 | -0.0129 | 0.0378 | -0.0248 | -0.0086 | 0.0162 | -0.0515 | -0.0164 | 0.0352 |
| | C-G-GCCF | -0.0575 | -0.0138 | 0.0437 | -0.0262 | -0.0091 | 0.0171 | -0.0538 | -0.0169 | 0.0369 |
| | NR-FHS | -0.0522 | -0.0134 | 0.0388 | -0.0260 | -0.0091 | 0.0170 | -0.0552 | -0.0178 | 0.0375 |
| | NR-H | -0.0509 | -0.0129 | 0.0380 | -0.0248 | -0.0087 | 0.0161 | -0.0520 | -0.0164 | 0.0356 |
| | NR-GCCF | -0.0617 | -0.0139 | 0.0478 | -0.0264 | -0.0093 | 0.0171 | -0.0551 | -0.0171 | 0.0380 |
| | H-L | | -0.0136 | 0.0136 | | -0.0084 | 0.0084 | | -0.0168 | 0.0168 |

| | | S&P500 | | | Euro/Dollar | | | IBEX35 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Model | | Lower Limit | Upper Limit | Width | Lower Limit | Upper Limit | Width | Lower Limit | Upper Limit | Width |
| GARCH-N | C-G-N | -0.0224 | -0.0204 | 0.0020 | -0.0167 | -0.0149 | 0.0018 | -0.0257 | -0.0234 | 0.0023 |
| | C-G-FHS | -0.0293 | -0.0219 | 0.0073 | -0.0198 | -0.0158 | 0.0040 | -0.0372 | -0.0280 | 0.0092 |
| | C-G-H | -0.0282 | -0.0221 | 0.0062 | -0.0205 | -0.0159 | 0.0046 | -0.0380 | -0.0281 | 0.0099 |
| | C-G-GCCF | -0.0285 | -0.0210 | 0.0075 | -0.0209 | -0.0166 | 0.0043 | -0.0330 | -0.0231 | 0.0099 |
| | NR-FHS | -0.0292 | -0.0216 | 0.0075 | -0.0197 | -0.0157 | 0.0040 | -0.0374 | -0.0275 | 0.0099 |
| | NR-H | -0.0277 | -0.0209 | 0.0069 | -0.0205 | -0.0152 | 0.0053 | -0.0379 | -0.0267 | 0.0112 |
| | NR-GCCF | -0.0286 | -0.0155 | 0.0131 | -0.0212 | -0.0160 | 0.0052 | -0.0328 | -0.0164 | 0.0165 |
| EGARCH-N | C-G-N | -0.0521 | -0.0135 | 0.0386 | -0.0275 | -0.0097 | 0.0178 | -0.0519 | -0.0168 | 0.0352 |
| | C-G-FHS | -0.0639 | -0.0160 | 0.0479 | -0.0306 | -0.0105 | 0.0201 | -0.0651 | -0.0207 | 0.0444 |
| | C-G-H | -0.0645 | -0.0160 | 0.0485 | -0.0312 | -0.0106 | 0.0206 | -0.0678 | -0.0214 | 0.0463 |
| | C-G-GCCF | -0.0623 | -0.0139 | 0.0484 | -0.0326 | -0.0109 | 0.0217 | -0.0664 | -0.0211 | 0.0453 |
| | NR-FHS | -0.0657 | -0.0160 | 0.0497 | -0.0308 | -0.0106 | 0.0202 | -0.0646 | -0.0203 | 0.0443 |
| | NR-H | -0.0642 | -0.0156 | 0.0486 | -0.0310 | -0.0104 | 0.0206 | -0.0672 | -0.0210 | 0.0462 |
| | NR-GCCF | -0.0573 | -0.0106 | 0.0466 | -0.0325 | -0.0113 | 0.0212 | -0.0658 | -0.0210 | 0.0448 |

# Chapter 5

# Summary of Conclusions and Future Research

Risk management has become very important because of important financial disasters that have caused big losses in banks and financial institutions. Due to that, risk managers have paid attention to the methods for quantifying the risk involved in its financial decisions. For this purpose, the Basel Committee establishes that the measure of risk financial institutions must use is the VaR. However, in the last years, alternative measures have came up which satisfy certain theoretical properties. The inconvenient about these measures is that some of them are very difficult to implement in practice. The ES is one of the most popular alternative measures for the VaR. Through this thesis we focus on different aspects related with these two measures: estimation methods, backtesting procedures, confidence level, forecast horizon and the uncertainty related to their estimation. Next, we describe the main contribution of the thesis.

In Chapter 2 we describe and illustrate alternative estimators of $VaR$ and $ES$. We show that the computational cost of both measures is similar. Therefore, if the regulatory agencies would decide to change the risk measure that must be used for financial institutions, the structure developed for the $VaR$ will still be useful for the estimation of the $ES$. By implementing alternative procedures to the same series of returns, we conclude that there are significant differences between the estimates obtained by alternative procedures and according to our results it is very important to choose an adequate specification of the conditional variance in order to have adequate estimates of the $VaR$ and $ES$.

In Chapter 3 we analyze if the conclusions about the chosen models change with the confidence level and horizon. The asymmetry involved in the specification of the conditional variance as well as the asymmetry in the distribution of the residuals, Bootstrap and $EVT$ with the $Hill$ estimator provide better $VaR$ forecasts regardless the confidence level. Furthermore, at least for the $S\&P500$ returns, the adequate specification of the variance is more important than the assumption for the error distribution for both measures of risk when assuming different confidence levels. On the other hand, with respect to the horizon, the results obtained for the one-step ahead $VaR$ and $ES$ can be generalized to the ten-steps ahead. Additionally, we conclude that there are differences when estimating the ten-steps ahead $VaR$ using daily or fortnightly data. However, one issue related to choosing fortnightly data is the availability of information, because, according to the Basel Committee, at least 250 observations must be used for estimation, which, in terms of fortnightly data means at least seven years. The same problem has to be faced for backtesting purposes because some times it is needed to backtest over a large number of observations.

In Chapter 4, we extend the bootstrap procedure proposed by Christoffersen and GonÇalves (2005) by taking into account the uncertainty associated to parameter estimation when constructing prediction intervals for the $VaR$ and the $ES$. We propose a second bootstrap step which avoid an extra source of uncertainty by bootstrapping directly from the original standardized residuals when computing the quantile of the error distribution. We also apply the smoothed iterated bootstrap method of Ho and Lee (2005). We observe that with the smoothed iterated method we can improve the coverage, but it is a very time consuming method making the computational time grow. By several Monte Carlo experiments we show that in most of the cases, our procedure produces coverages closer to the nominal for the $VaR$ and the $ES$.

Some of the directions for further research are:

- To analyze whether the importance of choosing an adequate specification for the conditional variance as opposite to choosing a correct distribution of the errors can be generalized to other time series of returns.

- The extension of the analysis carried out in Chapter 2 to incorporate Stochastic Volatility ($SV$) models in the comparison

- It would also interesting to incorporate in the comparison the estimates of the $VaR$ based on ultra-high-frequency data volatility measures.

- An interesting topic is the extension of the analysis carried out in Chapter 4 and to construct confidence intervals for the $VaR$ and the $ES$ when the evolution of the volatility is modeled by using Stochastic Volatility ($SV$) instead of $GARCH$ models

- It will also be interesting to know the behavior of the second bootstrap step procedure in the presence of misspecification.

# Bibliography


Acerbi, C., C. Nordio, and C. Sirtori (2001). Expected shortfall as a tool for financial risk management. *Working Paper Milan, AbaxBank*.

Acerbi, C. and D. Tasche (2002). Expected shortfall: a natural coherent alternative to value at risk. *Economic Notes 31*, 379–388.

Angelidis, T., A. Benos, and S. Degiannakis (2005). The use of GARCH models in VaR estimation. *Statistical Methodology 2*, 49–57.

Angelidis, T. and S. Degiannakis (2007). Backtesting VaR models: An expected shortfall approach. *Working Paper, University of Crete*.

Artzner, P., F. Delbaen, J. Eber, and D. Heath (1997). Thinking coherently. *Risk 10*, 68–71.

Artzner, P., F. Delbaen, J. Eber, and D. Heath (1999). Coherent measures of risk. *Mathematical Finance 2*, 203–228.

Bali, T. and P. Theodossiou (2007). A conditional-sgt-VaR approach with alternative garch models. *Annals of Operations Research 151*, 241–267.

Bams, D., T. Lehnert, and C. Wolff (2005). An evaluation framework for alternative var models. *Journal of International Money and Finance 24*, 944–958.

Bao, Y., T.-H. Lee, and B. Saltoglu (2006). Evaluating predictive performance of value-at-risk models in emerging markets: a reality check. *Journal of Forecasting 25*, 101 – 128.

Bao, Y. and A. Ullah (2004). Bias of value-at-risk model. *Finance Research Letters 1*, 241–249.

Barone-Adesi, G. and K. Giannopoulos (2001). Non-parametric VaR techniques. myths and realities. *Economic Notes 30*, 167–181.

Barone-Adesi, G., K. Giannopoulos, and L. Vosper (1999). VaR without correlations for portfolios of derivative securities. *The Journal of Future Markets 19*, 583–602.

Basak, S. and A. Shapiro (2001). Value-at-risk based risk management: Optimal policies and asset prices. *Review of Financial Studies 14*, 371–405.

Basel Committee on Banking Supervision (1996a). Amendment to the capital accord to incorporate market risks, bank for international settlements. Technical report.

Basel Committee on Banking Supervision (1996b). Supervisory framework for the use of backtesting in conjunction with the internal models approach to market risk capital requirements, bank for international settlements. Technical report.

Berkowitz, J. (2001). Testing density forecast, with applications to risk management. *Journal of Business and Economic Statistics 19*, 465–474.

Berkowitz, J., P. Christoffersen, and D. Pelletier (2006). Evaluating value-at-risk models with desk-level data. *Working Paper Series 010 Department of Economics, North Carolina State University*.

Berkowitz, J. and L. Kilian (2000). Recent developments in bootstrapping time series. *Econometric Review 19*, 1–48.

Billio, M. and L. Pelizzon (2000). Value-at-risk: a multivariate switching regime approach. *Journal of Empirical Finance 7*, 531–554.

Billio, M. and D. Sartore (2003). *Stochastic Volatility Models: A Survey with Applications to Option Pricing and Value at Risk*. West Sussex, UK: John Wiley.

Black, F. (1976). Studies of stock price volatility changes. *Proceedings of the Business and Economic Statistics Section*, 177–181.

Blake, D., C. A., and K. Dowd (2000). Extrapolating var by the square-root rule. *Financial Engineering News 3*, 7.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*, 307–327.

Brownless, C. T. and G. M. Gallo (2010). Comparison of volatility measures: a risk management perspective. *Journal of Financial Econometrics 8*, 29.

Cai, Z. (2002). Regression quantiles for time series data. *Econometric Theory 18*, 169–192.

Cai, Z. and X. Wang (2008). Nonparametric estimation of conditional VaR and expected shortfall. *Journal of Econometrics 147*, 120–130.

Chan, K. and P. Gray (2009). Using extreme value theory to measure value-at-risk for daily electricity spot prices. *Journal of Forecasting 22*, 283–300.

Chan, N., S. Deng, L. Peng, and Z. Xia (2007). Interval estimation for the conditional value at risk based on GARCH models with heavy tailed innovations. *Journal of Econometrics 137*, 556–576.

Chen, M. Y. and J. Chen (2005). Application of quantile regression to estimation of value at risk. *Review of Financial Risk Management 1*, 1–15.

Chen, S. and C. Tang (2005a). Nonparametric inference of value-at-risk for dependent financial returns. *Journal of Financial Econometrics 3*, 227–255.

Chen, S. and C. Tang (2005b). Nonparametric inference of value at risk for dependent financial returns. *Journal of Financial Econometrics 3*, 227–255.

Chou, J., H. Yu, and Z. Chen (2008). Interval estimation of value at risk for Taiwan weighted stock index based on extreme value theory. *Journal of the Chinese Institute of Industrial Engineers 25*, 31–42.

Christoffersen, P. (1998). Evaluating interval forecast. *Journal of Business and Economic Statistics 39*, 841–862.

Christoffersen, P. and F. Diebold (2000). How relevant is volatility forecasting for financial risk management? *The Review of Economics and Statistics 82*, 12–22.

Christoffersen, P. and S. GonÇalves (2005). Estimation risk in financial risk management. *Journal of Risk 7*, 1–28.

Christoffersen, P., J. Hahn, and A. Inoue (2001). Testing and comparing value-at-risk measures. *Journal of Empirical Finance 8*, 325–342.

Danielsson, J., L. de Haan, L. Peng, and C. G. de Vries (2001). Using a bootstrap method to choose the sample fraction in tail index estimation. *Journal of multivariate analysis 76*, 226–248.

Danielsson, J. and C. de Vries (1997). Tail index and quantile estimation with very high frequency data. *Journal of Empirical Finance 4*, 241–257.

Danielsson, J. and C. de Vries (2000). Value at risk and extreme returns. *Annales d'Economie et Statistique 60*, 239–270.

Danielsson, J. and C. Hartmann, P. De Vries (1998). The cost of conservatism: Extreme returns, value-at risk, and the basle multiplicaiton factor. *RISK 11*, 101–103.

Danielsson, J., B. Jorgensen, M. Sarma, and C. G. de Vries (2006). Comparing downside risk measures for heavy tailed distributions. *Economics letters 92*, 202–208.

Danielsson, J., B. Jorgensen, M. Sarma, G. Samorodnitsky, and C. G. de Vries (2005). Subadditivity re-examined: The case for value-at-risk. *Discussion paper 549, Financial Markets Group, LSE*.

Danielsson, J. and J. Zigrand (2006). On time-scaling of risk and the square-root-of-time rule. *Journal of Banking and Finance 30*, 2701–2713.

DeRossi, G. and A. Harvey (2006). Time-varying quantiles. *University of Cambridge, Faculty of Economics, CWPE 0649*.

Diebold, F. (1988). *Empirical Modeling of Exchange Rate Dynamics.* New York: Springer-Verlag.

Diebold, F., A. Hickman, A. Inoue, and T. Schuermann (1997). Converting 1-day volatility to h-day volatitlity: Scaling by root-h is worse than you think. *Working Paper, Center for Financial Institutions, University of Pennsylvania*.

Diebold, F. X. and R. Mariano (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics 13*, 253–263.

Ding, Z., C. Granger, and R. Engle (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance 1*, 83–106.

Dowd, K. (2001). Estimating VaR with order statistics. *Journal of Derivatives 8*, 23–30.

Duffie, D. and J. Pan (1997). An overview of VaR. *Journal of Derivatives 4*, 7–49.

Eberlein, E., J. Kallsen, and J. Kristen (2003). Risk management based on stochastic volatility. *Journal of Risk 5*, 19–44.

Efron, B. (1981). Nonparametric standard errors and confidence intervals. *The Canadian Journal of Statistics*, 139–172.

Efron, B. (1982). The jackknife, the bootstrap, and other resampling plans. *Regional Conference Series in Applied Mathematics*, No.38. Philadelphia: SIAM.

Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American Statistical Association*, 171–185.

Embrechts, P., C. Klüppelberg, and T. Mikosch (2000). *Modelling Extremal Events for Insurance and Finance*. Londres: Springer.

Engle, R. (2003). Risk and volatility: Econometric models and financial practice. *Nobel Lecture*.

Engle, R. and S. Manganelli (2004). Caviar: Conditional autoregressive value at risk by regression quantiles. *Journal of Business and Economic Statistics 22*, 367–381.

Escanciano, J. and J. Olmo (2010). Backtesting parametric value at risk with estimation risk. *Journal of Business and Economic Statistics 28*, 36–51.

Fabozzi, F. and R. Tunaru (2006). On risk management problems related to a coherence property. *Quantitative Finance 6*, 75–81.

Fan, J. (1992). Design adaptive nonparametric regression. *Journal of the American Statististical Association 87*, 998–1004.

Fan, J. and J. Gu (2003). Semiparametric estimation of value at risk. *The Econometrics Journal 6*, 261–290.

GenÇay, R. and F. SelÇuk (2004). Extreme value theory and value at risk: relative performance in emerging markets. *International Journal of Forecasting 20*, 287–303.

Giacomini, R. and I. Komunjer (2005). Evaluation and combination of conditional quantile forecasts. *Journal of Business & Economic Statistics 23*, 416–431.

Giacomini, R. and H. White (2006). Tests of conditional predictive ability. *Econometrica 74*, 1545–1578.

Giamouridis, D. (2006). Estimation risk in financial risk management: A correction. *Journal of Risk 8*, 121–125.

Giannopoulos, K. (2003). VaR modelling on long run horizons. *Automation and Remote Control 64*, 1094–1100.

Giannopoulos, K. and R. Tunaru (2005). Coherent risk measures under filtered historical simulation. *Journal of Banking and Finance 29*, 979–996.

Gilli, M. and E. Kllezi (2006). An application of extreme value theory for measuring financial risk. *Computational Economics 27*, 207–228.

Giot, P. and S. Laurent (2003). Value-at-risk for long and short trading positions. *Journal of Applied Econometrics 18*, 641–663.

Glosten, L., R. Jagannathan, and D. Runkle (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance 48*, 1779–1801.

Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Annals of Mathematics 44*, 423–453.

Gonzalo, J. and J. Olmo (2004). Which extreme values are really extreme? *Journal of Financial Econometrics 2*, 349–369.

Gourieroux, C. and J. Jasiak (2006). Dynamic quantile models. *York University, Department of Economics Working Paper*.

Granger, C. W. (2002). Some comments on risk. *Journal of Applied Econometrics 17*, 447–456.

Haas, M. and A. Kondratyev (2000). Value at risk and expected shortfall with confidence bands: an extreme value theory approach. *Working Paper, Research Center CAESAR*.

Hall, P. and Q. Yao (2003). Data tilting for time series. *Journal of the Royal Statistical Society, Series B 65*, 425–442.

Hansen, B. (1994). Autoregressive conditional density estimation. *International Economic Review 35*, 705–730.

Hansen, P. (2005). A test for superior predictive ability. *Journal of Business and Economic Statistics 23*, 365–380.

Harmantzis, F., L. Miao, and Y. Chien (2006). Empirical study of value-at-risk and expected shortfall models with heavy tails. *The Journal of Risk Finance Incorporating Balance Sheet 7*, 117–135.

Hartz, C., M. S. and M. Paolella (2006). Accurate value-at-risk forecasting based on the normal-GARCH model. *Computational Statistics & Data Analysis 51*, 2295–2312.

He, C., A. Silvennoinen, and T. Teräsvirta (2008). Parameterizing unconditional skewness in models for financial time series. *Journal of Financial Econometrics 6*, 208–230.

Hendricks, D. (1996). Evaluation of value at risk models using historical data. *Federal Reserve Bank of New York Economic Policy Review, April*, 39–69.

Hentschel, L. (1995). All in the family: Nesting symmetric and asymmetric GARCH models. *Journal of Financial Economics 39*, 71–104.

Hill, B. (1975). A simple general approach to inference about the tail of a distribution. *Annals of Statistics 3*, 1163–1174.

Ho, Y. and S. Lee (2005). Iterated smoothed bootstrap confidence intervals for population quantiles. *The Annals of statistics*, 437–462.

Hull, J. and A. White (1998). Incorporating volatility updating into the historical simulation method for value at risk. *Journal of Risk 1*, 5–19.

Inui, K. and M. Kijima (2005). On the significance of expected shortfall as a coherent risk measure. *Journal of Banking and Finance 29*, 853–864.

Jalal, A. and M. Rockinger (2008). Predicting tail-related risk measures: The consequences of using GARCH filters for non-GARCH data. *Journal of Empirical Finance 15*, 868–877.

Jorion, P. (1990). *Value at Risk: the new benchmark for controlling market risk.* University of California: McGraw-Hill.

Jorion, P. (1997). In defense of VaR. *Derivatives Strategy 2*, 20–23.

Ju, X. and N. D. Pearson (1999). Using value-at-risk to control risk taking: How wrong can you be? *Journal of Risk 1*, 5–36.

Kerkhof, J. and B. Melenberg (2002). Backtesting for risk-based regulatory capital. *Discussion Paper, Tilburg University, Center for Economic Research*.

Koenker, R. and G. Bassett (1978). Regression quantiles. *Econometrica 46*, 33–50.

Krause, A. (2003). Exploring the limitations of value at risk: How good is it in practice? *Journal of Risk Finance 4*, 19–28.

Kuester, K., S. Mittnik, and M. Paolella (2006). Value-at-risk prediction: A comparison of alternative strategies. *Journal of Financial Econometrics 4*, 53–89.

Kupiec, P. (1995). Techniques for verifying the accuracy of risk management models. *Journal of Derivatives 3*, 73–84.

Lambert, P. and S. Laurent (2000). Modelling skewness dynamics in series of financial data. *Discussion Paper, Institute de Statistique, Louvain-la-Neuve*.

Lan, H., B. Nelson, and J. Staum (2008). A confidence interval procedure for expected shortfall risk measurement via two-level simulation. *Working Paper, Department of Industrial Engineering and Management Sciences, Northwestern University*.

Levy, H. (1992). Stochastic dominance expected utility: survey and analysis. *Management Science 38*, 555–593.

Lima, R. and B. Néri (2007). Comparing value-at-risk methodologies. *Brazilian Review of Econometrics 27*, 1–25.

Longerstaey, J. and L. More (1995). *Introduction to RiskMetrics* (4a edición ed.). New York: Morgan Guaranty Trust Company.

Longin, F. (2001). Beyond the VaR. *Journal of Derivatives 8*, 36–48.

Lopez, J. (1999). Methods for evaluating value-at-risk estimates. *Federal Reserve Bank of San Francisco Economic Review 2*, 3–17.

Luce, R. (1981). Corrections to 'some possible measures of risk'. *Theory and Decision 13*, 381.

Machina, M. and M. Rothschild (1987). *Risk*. London: Macmillan.

Manganelli, S. and R. Engle (2001). Value at risk models in finance. *Working Paper, European Central bank*.

Martins-Filho, C. and F. Yao (2006). Estimation of value-at-risk and expected shortfall based on nonlinear models of return dynamics and extreme value theory. *Studies in Nonlinear Dynamics & Econometrics 10*, 107–149.

McAleer, M. (2009). The ten commandments for optimizing value-at-risk and daily capital charges. *Forthcoming in Journal of Economic Surveys*.

McAleer, M. and B. da Veiga (2008). Single index and portfolio models for forecasting value-at-risk thresholds. *Journal of Forecasting 27*, 217–235.

McLeod, A. and W. Li (1983). Diagnostic checking arma time series models using squared-residual autocorrelations. *Journal of the Time Series Analysis 4*, 269–273.

McNeil, A. J. and R. Frey (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of Empirical Finance 7*, 271–300.

Mittnik, S. and M. Paolella (2000). Conditional density and value-at-risk prediction of asian currency exchange rates. *Journal of Forecasting 19*, 313–333.

Morgan, J. (1995). Riskmetrics. *RiskMetrics Technical Document. 3a edición*.

Nelson, D. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica 59*, 347–370.

Newey, W. and K. West (1987). A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica 55*, 703–708.

Nieto, M. and E. Ruiz (2009). Measuring financial risk: Comparison of alternative procedures to estimate VaR and ES. *Working Paper 08-73, Universidad Carlos III de Madrid*.

Nystrom, K. and J. Skoglund (2002). Univariate extreme value theory, GARCH and measures of risk. *Preprint, Swedbank*.

Pascual, L., J. Romo, and E. Ruiz (2006). Bootstrap prediction for returns and volatilities in GARCH models. *Computational Statistics and Data Analysis 50*, 2293–2312.

Peng, L. and Y. Qi (2003). Confidence intervals for high quantiles of a heavy tailed distribution. *Technical Report*, Georgia Institute of Technology, United States.

Politis, D. and J. Romano (1994). The stationary bootstrap. *Journal of American Statistical Association 89*, 1303–1313.

Politis, D. and H. White (2004). Automatic block-length selection for the dependent bootstrap. *Econometric Review 23*, 53–70.

Pownall, R. and K. Koedijk (1999). Capturing downside risk in financial markets: the case of the asian crisis. *Journal of International Money and Finance 18*, 853–870.

Rodriguez, J. and E. Ruiz (2005). A powerful test for conditional heteroscedasticity for financial time series with highly persistent volatilities. *Statistica Sinica 15*, 505–525.

Rodríguez, M. and E. Ruiz (2009). GARCH models with leverage effect: differences and similarities. *WP 09-03(02), Universidad Carlos III de Madrid*.

Ruiz, E. and L. Pascual (2002a). Bootstraping financial time series. *Journal of Economic Surveys 16*, 271–300.

Ruiz, E. and L. Pascual (2002b). Bootstrapping financial time series. *Journal of Economic Surveys 16*, 271–300.

Sadorsky, P. (2005). Stochastic volatility forecasting and risk management. *Applied Financial Economics 15*, 121–135.

Santos, A., F. Nogales, and E. Ruiz (2009). Comparing univariate and multivariate models to forecast portfolio value-at-risk. *WP 09-72, Universidad Carlos III de Madrid*.

Sarma, M., S. Thomas, and A. Shah (2003). Selection of value at risk models. *Journal of Forecasting 22*, 337–358.

Szegö, G. (2002). Measures of risk. *Jornal of Banking and Finance 26*, 1253–1272.

Theodossiou, P. (1998). Financial data and the skewed generalized t distribution. *Management Science 44*, 1650–1661.

Tobin, J. (1969). Comment on borch and feldstein. *The Review of Economic Studies 36*, 13–14.

Tsiang, S. (1972). The rationale of the mean-stardard deviation analysis, skewness preference, and the demand for money. *The American Economic Review 62*, 354–371.

Vlaar, P. J. G. (2000). Value at risk models for dutch bond portfolios. *Journal of Banking and Finance 24*, 1131–1154.

Wang, S. (1998). An actuarial index of the right-tail index. *North American Actuarial Journal 2*, 88–101.

Wang, S., V. Young, and H. Panjer (1997). Axiomatic characterization of insurance prices. *Insurance: Mathematics and Economics 25*, 173–183.

Wong, W. (2010). Backtesting value-at-risk based on tail losses. *Journal of Empirical Finance 17*, 526–538.

Yamai, Y. and T. Yoshiba (2005). Value-at-risk versus expected shortfall: A practical perspective. *Journal of Banking & Finance 29*, 997–1015.

Yu, K. and M. Jones (1998). Local linear quantile regression. *Journal of American Statistical Association 93*, 228–237.

Zakoian, J.-M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control 18*, 931–995.